

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: April 15, 2005, 14:09:02 ; Search time 137 Seconds  
(without alignments)

713.265 Million cell updates/sec

Title: US-10-809-655-9

Perfect score: 1528  
Sequence: 1 MSPRGTGCSAGILMTVGMIL.....LDQLPTMPGDDALSEWNW 294

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1528	100.0	294	US-10-181-642-9	Sequence 9, Appli
2	1528	100.0	294	US-10-809-654-9	Sequence 9, Appli
3	1528	100.0	294	US-10-809-655-9	Sequence 9, Appli
4	762.5	49.9	163	US-10-411-224-105	Sequence 105, App
5	762.5	49.9	163	US-10-047-021-105	Sequence 105, App
6	593	38.8	124	US-10-047-021-119	Sequence 119, App
7	593	38.8	124	US-10-411-224-119	Sequence 119, App
8	110	7.2	1315	US-10-369-493-3346	Sequence 3346, Ap
9	103	6.7	255	US-09-833-245-1568	Sequence 1568, Ap
10	100	6.5	604	US-10-437-963-174055	Sequence 174055, Ap
11	99	6.5	255	US-09-833-245-1567	Sequence 1567, App
12	99	6.5	255	US-10-052-586-538	Sequence 538, App
13	99	6.5	255	US-10-174-590-538	Sequence 538, App

14	99	6.5	255	US-10-176-758-538	Sequence 538, App
15	99	6.5	255	US-10-175-737-538	Sequence 538, App
16	99	6.5	255	US-10-174-581-538	Sequence 538, App
17	99	6.5	255	US-10-176-483-538	Sequence 538, App
18	99	6.5	255	US-10-176-749-538	Sequence 538, App
19	99	6.5	255	US-10-176-914-538	Sequence 538, App
20	99	6.5	255	US-10-176-915-538	Sequence 538, App
21	99	6.5	255	US-10-173-706-538	Sequence 538, App
22	99	6.5	255	US-10-175-728-538	Sequence 538, App
23	99	6.5	255	US-10-175-752-538	Sequence 538, App
24	99	6.5	255	US-10-176-482-538	Sequence 538, App
25	99	6.5	255	US-10-176-757-538	Sequence 538, App
26	99	6.5	255	US-10-176-913-538	Sequence 538, App
27	99	6.5	255	US-10-180-552-538	Sequence 538, App
28	99	6.5	255	US-10-180-557-538	Sequence 538, App
29	99	6.5	255	US-10-173-700-538	Sequence 538, App
30	99	6.5	255	US-10-174-572-538	Sequence 538, App
31	99	6.5	255	US-10-174-579-538	Sequence 538, App
32	99	6.5	255	US-10-174-582-538	Sequence 538, App
33	99	6.5	255	US-10-174-588-538	Sequence 538, App
34	99	6.5	255	US-10-175-739-538	Sequence 538, App
35	99	6.5	255	US-10-175-740-538	Sequence 538, App
36	99	6.5	255	US-10-175-743-538	Sequence 538, App
37	99	6.5	255	US-10-176-492-538	Sequence 538, App
38	99	6.5	255	US-10-176-492-538	Sequence 538, App
39	99	6.5	255	US-10-176-747-538	Sequence 538, App
40	99	6.5	255	US-10-176-750-538	Sequence 538, App
41	99	6.5	255	US-10-176-985-538	Sequence 538, App
42	99	6.5	255	US-10-176-987-538	Sequence 538, App
43	99	6.5	255	US-10-176-992-538	Sequence 538, App
44	99	6.5	255	US-10-176-993-538	Sequence 538, App
45	99	6.5	255	US-10-184-658-538	Sequence 538, App

#### ALIGNMENTS

RESULT 1  
US-10-181-642-9  
Sequence 9, Application US/10181642  
Publication No. US20030087375A1  
GENERAL INFORMATION:  
APPLICANT: Herr, John C.  
APPLICANT: Shetty, Jagathapala  
APPLICANT: Wolkowicz, Michael  
APPLICANT: Hayes, Friederike  
APPLICANT: Hao, Zhonglin  
TITLE OF INVENTION: Sperm Specific Proteins  
FILE REFERENCE: 00497-02  
CURRENT APPLICATION NUMBER: US/10/181,642  
PRIOR APPLICATION NUMBER: 60/176,885  
PRIORITY DATE: 2000-01-19  
PRIORITY DATE: 2000-01-19  
NUMBER OF SEQ. IDS: 19  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 9  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-181-642-9

Query Match 100.0%; Score 1528; DB 14; Length 294;  
Best Local Similarity 100.0%; Pred. No. 5e-150;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPRGTGCSAGILMTVGMILAGLQSGRTVTAQVADAGLAHGBGEBETENDSEYAE 60  
DB 1 MSPRGTGCSAGILMTVGMILAGLQSGRTVTAQVADAGLAHGBGEBETENDSEYAE 60  
QY 61 NYAPPEPTEVSNRNVVAVKVEFGMCTVTCGIVRAVILLNCGPGBSKCVAVYEBRGPTD 120  
DB 61 NYAPPEPTEVSNRNVVAVKVEFGMCTVTCGIVRAVILLNCGPGBSKCVAVYEBRGPTD 120

QY 121 CGMKPISSESVRLACIHTSPINRFKYMMLKRODOOSIILVNDAILVRKESHPLA 180  
DB 121 CGMKPISSESVRLACIHTSPINRFKYMMLKRODOOSIILVNDAILVRKESHPLA 180  
QY 181 FECDLNNNEIVATIKFTVYTSSELQMRSSLPATDAILFVLITGVIIICVFIIFLLIFI 240  
DB 181 FECDLNNNEIVATIKFTVYTSSELQMRSSLPATDAILFVLITGVIIICVFIIFLLIFI 240  
QY 241 IINAAVAKAFWGAKASTPEVQSSSVRYKOSTSLDOLPTMPGDDALSEWNE 294  
DB 241 IINAAVAKAFWGAKASTPEVQSSSVRYKOSTSLDOLPTMPGDDALSEWNE 294

## RESULT 2

US-10-809-654-9  
Sequence 9, Application US/10809654  
Publication No. US20040161824A1  
GENERAL INFORMATION:  
APPLICANT: Herr, John C.  
APPLICANT: Shetty, Jagathapala  
APPLICANT: Wolkowicz, Michael  
APPLICANT: Jayes, Friederike  
APPLICANT: Hao, Zhonglin  
TITLE OF INVENTION: Sperm Specific Proteins  
FILE REFERENCE: 00497-02  
CURRENT APPLICATION NUMBER: US/10/809,654  
CURRENT FILING DATE: 2004-03-25  
PRIOR APPLICATION NUMBER: 60/076,885  
PRIOR FILING DATE: 2000-01-19  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-809-654-9

Query Match 100.0%; Score 1528; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 5e-150; Indels 0; Gaps 0;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSPRGTCGAGLMTVGMILLAGLSARGNTVTAADVADAGLAHGEGBEETENDSETAE 60  
DB 1 MSPRGTCGAGLMTVGMILLAGLSARGNTVTAADVADAGLAHGEGBEETENDSETAE 60  
QY 61 NYAPETEDVSNRNVKVEFGMCTVTCGIGREVILLTNGCGGSKCVARVEECRGPTD 120  
DB 61 NYAPETEDVSNRNVKVEFGMCTVTCGIGREVILLTNGCGGSKCVARVEECRGPTD 120  
QY 121 CGMKPISSESVRLACIHTSPINRFKYMMLKRODOOSIILVNDAILVRKESHPLA 180  
DB 121 CGMKPISSESVRLACIHTSPINRFKYMMLKRODOOSIILVNDAILVRKESHPLA 180  
QY 181 FECDLNNNEIVATIKFTVYTSSELQMRSSLPATDAILFVLITGVIIICVFIIFLLIFI 240  
DB 181 FECDLNNNEIVATIKFTVYTSSELQMRSSLPATDAILFVLITGVIIICVFIIFLLIFI 240  
QY 241 IINAAVAKAFWGAKASTPEVQSSSVRYKOSTSLDOLPTMPGDDALSEWNE 294  
DB 241 IINAAVAKAFWGAKASTPEVQSSSVRYKOSTSLDOLPTMPGDDALSEWNE 294

## RESULT 3

US-10-809-655-9  
Sequence 9, Application US/10809655  
Publication No. US20040161825A1  
GENERAL INFORMATION:  
APPLICANT: Herr, John C.  
APPLICANT: Shetty, Jagathapala  
APPLICANT: Wolkowicz, Michael  
APPLICANT: Jayes, Friederike  
APPLICANT: Hao, Zhonglin  
TITLE OF INVENTION: Sperm Specific Proteins

FILE REFERENCE: 00497-02  
CURRENT APPLICATION NUMBER: US/10/809,655-9  
CURRENT FILING DATE: 2004-03-25  
PRIOR APPLICATION NUMBER: 60/76,885  
PRIOR FILING DATE: 2000-01-19  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 294  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-809-655-9

Query Match 100.0%; Score 1528; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 5e-150; Indels 0; Gaps 0;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSPRGTCGAGLMTVGMILLAGLSARGNTVTAADVADAGLAHGEGBEETENDSETAE 60  
DB 1 MSPRGTCGAGLMTVGMILLAGLSARGNTVTAADVADAGLAHGEGBEETENDSETAE 60  
QY 61 NYAPETEDVSNRNVKVEFGMCTVTCGIGREVILLTNGCGGSKCVARVEECRGPTD 120  
DB 61 NYAPETEDVSNRNVKVEFGMCTVTCGIGREVILLTNGCGGSKCVARVEECRGPTD 120  
QY 121 CGMKPISSESVRLACIHTSPINRFKYMMLKRODOOSIILVNDAILVRKESHPLA 180  
DB 121 CGMKPISSESVRLACIHTSPINRFKYMMLKRODOOSIILVNDAILVRKESHPLA 180  
QY 181 FECDLNNNEIVATIKFTVYTSSELQMRSSLPATDAILFVLITGVIIICVFIIFLLIFI 240  
DB 181 FECDLNNNEIVATIKFTVYTSSELQMRSSLPATDAILFVLITGVIIICVFIIFLLIFI 240  
QY 241 IINAAVAKAFWGAKASTPEVQSSSVRYKOSTSLDOLPTMPGDDALSEWNE 294  
DB 241 IINAAVAKAFWGAKASTPEVQSSSVRYKOSTSLDOLPTMPGDDALSEWNE 294

## RESULT 4

US-10-411-224-105  
Sequence 105, Application US/10411224  
Publication No. US20030166906A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 50 Human Secreted Proteins  
FILE REFERENCE: P2016P1  
CURRENT APPLICATION NUMBER: US/10/411,224  
CURRENT FILING DATE: 2003-04-11  
PRIOR APPLICATION NUMBER: US/09/722,329  
PRIOR FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: 09/262,109  
PRIOR FILING DATE: 1999-03-04  
PRIOR APPLICATION NUMBER: 60/057,626  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: 60/057,663  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: 60/057,669  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: 60/058,667  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,974  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,973  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,666  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/090,112  
PRIOR FILING DATE: 1998-06-22  
NUMBER OF SEQ ID NOS: 206  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 105  
LENGTH: 163  
TYPE: PRT

```

; LOCATION: (113)
; OTHER INFORMATION: Xaa equals any amino acid
US-10-047-021-105

Query Match      49.9%; Score 762.5; DB 15; Length 163;
Best Local Similarity 94.2%; Pred. No. 9e-71;
Matches 147; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MSPRGTCAGLMTYGMILLAGLQSGARGTNTVAAYVDAGLAHGEGBEETENNDSSTAE 60
DB 1 MSPRGTCAGLMTYGMILLAGLQSGARGTNTVAAYVDAGLAHGEGBEETENNDSSTAE 60

QY 61 NYAPPEDEDVSNRNVAKEVEFGMCTVTCIGVREVIITNGCPGGESEKCVYRVEBCRGPTD 120
DB 61 NYAPPEDEDVSNRNVAKEVEFGMCTVTCIGVREVIITNGCPGGESEKCVYRVEBCRGPTD 120

QY 121 CGMGKPISESLSESVRLACHTSPPLNRPKNNMKLLRQ 156
DB 121 CGMGKPISESLSESVRLACHTSPPLNRPKNNMKLLRQ 155

RESULT 6
US-10-047-021-119
; Sequence 119, Application US/10047021
; Publication No. US20040002591A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: P2016P2
; CURRENT APPLICATION NUMBER: US/10/047,021
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/262,066
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 09/722,329
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/262,109
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: PCT/US98/18360
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: US 60/057,626
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,663
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/057,669
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/058,667
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,974
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,973
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/058,666
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: US 60/090,112
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 119
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals any amino acid
US-10-047-021-119

Query Match      38.8%; Score 593; DB 15; Length 124;
Best Local Similarity 98.3%; Pred. No. 2.6e-53;
Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSPRGTCAGLMTYGMILLAGLQSGARGTNTVAAYVDAGLAHGEGBEETENNDSSTAE 60

```

Db 1 MSPRGTGCSAGLMTVGMILLAGLQSGARGTNTAAVQDAGLAHGESEETENDSETAE 60  
 QY 61 NYAPPEDEDVSNRNKVKVEFEFGMCTVTGIGVREVILTNCGPGSGSKCVRVREEC 115  
 Db 61 NYAPSETEDVSNRNKVKVEFEFGMCTVTGIGVREVILTNCGPGSGSKCVRVREEC 115

## RESULT 7

US-10-411-224-119  
 Sequence 119, Application US/10411224  
 Publication No. US20030166906A1

## GENERAL INFORMATION:

APPLICANT: Rosen et al.  
 TITLE OF INVENTION: 50 Human Secreted Proteins  
 FILE REFERENCE: P2016P1  
 CURRENT FILING DATE: 2003-04-11  
 PRIOR APPLICATION NUMBER: US/09/722,329  
 PRIOR FILING DATE: 2000-11-28  
 PRIOR APPLICATION NUMBER: 09/262,109  
 PRIOR FILING DATE: 1999-03-04  
 PRIOR APPLICATION NUMBER: 60/057,626  
 PRIOR FILING DATE: 1997-09-05  
 PRIOR APPLICATION NUMBER: 60/057,663  
 PRIOR FILING DATE: 1997-09-05  
 PRIOR APPLICATION NUMBER: 60/057,669  
 PRIOR FILING DATE: 1997-09-05  
 PRIOR APPLICATION NUMBER: 60/058,667  
 PRIOR FILING DATE: 1997-09-12  
 PRIOR APPLICATION NUMBER: 60/058,974  
 PRIOR FILING DATE: 1997-09-12  
 PRIOR APPLICATION NUMBER: 60/058,973  
 PRIOR FILING DATE: 1997-09-12  
 PRIOR APPLICATION NUMBER: 60/058,666  
 PRIOR FILING DATE: 1997-09-12  
 PRIOR APPLICATION NUMBER: 60/090,112  
 PRIOR FILING DATE: 1998-06-22  
 NUMBER OF SEQ ID NOS: 206  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 119  
 LENGTH: 125  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE: 1  
 NAME/KEY: SITE  
 LOCATION: (75)  
 OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
 FEATURE: 2  
 NAME/KEY: SITE  
 LOCATION: (125)  
 OTHER INFORMATION: Xaa equals stop translation  
 US-10-411-224-119

Query Match 38.8%; Score 593; DB 14; Length 125;  
 Best Local Similarity 98.3%; Pred. No. 2.6e-53;  
 Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSPRGTGCSAGLMTVGMILLAGLQSGARGTNTAAVQDAGLAHGESEETENDSETAE 60  
 Db 1 MSPRGTGCSAGLMTVGMILLAGLQSGARGTNTAAVQDAGLAHGESEETENDSETAE 60  
 QY 61 NYAPPEDEDVSNRNKVKVEFEFGMCTVTGIGVREVILTNCGPGSGSKCVRVREEC 115  
 Db 61 NYAPSETEDVSNRNKVKVEFEFGMCTVTGIGVREVILTNCGPGSGSKCVRVREEC 115

## RESULT 8

US-10-369-493-3346  
 Sequence 3346, Application US/10369493  
 Publication No. US20030233675A1  
 GENERAL INFORMATION:  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.  
 APPLICANT: Goldman, Barry S.

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 FILE REFERENCE: 38-10(52052)B  
 CURRENT FILING DATE: 2003-02-28  
 PRIOR APPLICATION NUMBER: US/10/369,493  
 PRIOR FILING DATE: 2002-02-21  
 NUMBER OF SEQ ID NOS: 47374  
 SEQ ID NO 3346  
 LENGTH: 1315  
 TYPE: PRT  
 ORGANISM: Neurospora crassa  
 US-10-369-493-3346

Query Match 7.2%; Score 110; DB 15; Length 1315;  
 Best Local Similarity 25.2%; Pred. No. 0.14;  
 Matches 61; Conservative 38; Mismatches 93; Indels 50; Gaps 14;

QY 41 LAHGESEETENDSETAENYA---PPEEDV---SNRNKVKVEFEFGMCTVTGIGVRE 94  
 Db 378 LAHESAANKETEBOVPKESMLLDHVDDDVGESESKNKTSESN----- 425  
 QY 95 VILTNCGPGSGS-KCVRVRECRGPTDGCW-----GKPISESL---ESVRLACIHTSP--- 143  
 Db 426 ---NNGAESSSGSEFTITDQDSEID--WGLRLASPLSOLSTELVRLKVTAGKPKDKR 480  
 QY 144 LNRFKYMKLL--KDDQOSILVNDSDALEVRKSHPLAFECOTLDNNEIYA---TIKF 197  
 Db 481 VNEFVGTLELPSQDQVWGAANVPREGDVK--AAPLSID-NPAMANTVIAASNATTAV 537  
 QY 198 TVYTSSELQMRSSIPATDALIF---VLTGVLICVITIFLLFIINMAAVKAFWGA 254  
 Db 538 IVYTGPOTRSAISTSPRSKKTGLLEYEINSLTKLC---ELTFLSLVALLEGFSTAK 593  
 QY 255 AS 256  
 Db 594 GN 595

## RESULT 9

US-09-833-245-1568  
 Sequence 1568, Application US/09833245  
 Publication No. US20040010134A1

## GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.  
 TITLE OF INVENTION: Albumin Fusion Proteins  
 FILE REFERENCE: PF546PCT  
 CURRENT FILING DATE: 2001-04-12  
 PRIOR APPLICATION NUMBER: 60/229,358  
 PRIOR FILING DATE: 2000-04-12  
 PRIOR APPLICATION NUMBER: 60/256,931  
 PRIOR FILING DATE: 2000-12-21  
 PRIOR APPLICATION NUMBER: 60/199,384  
 PRIOR FILING DATE: 2000-04-25  
 NUMBER OF SEQ ID NOS: 2267  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 1568  
 LENGTH: 255  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-833-245-1568

Query Match 6.7%; Score 103; DB 11; Length 255;  
 Best Local Similarity 24.8%; Pred. No. 0.069;  
 Matches 41; Conservative 21; Mismatches 65; Indels 38; Gaps 8;

QY 58 TAENYAPPEDEDVSNRNKVKVEFEFGMCTVTGIGVREVILTNCGPGSGSKCVRVREEC 116  
 Db 25 TPKTLAIPKLOBAVGKYL--INATTCVTGIGVKEFEVCEVGPDPGRKRCOTRLBCL 82

```

QY 117 GPTGCG-----GKPTSLSESVRLAIIHNSPL---NRFFYMKRLRPOQSIILVN 165
Db 83 TMTICGMHFTLLICK-----EPFLSTLSIDILHPGGEAFPTFTRLAR---GVISID 131
QY 166 DSALLFRKSHSLPLAF-----CDT--LDNNELVATIKF 197
Db 132 DEVRKPFQANSHFVFKYTAQOEDSGTRCDVOLVKNLRILVKRIYF 176

```

RESULT 10  
US-10-437-963-174055

Query Match	6.5%	Score 100;	DB 16;	length 604;
Best Local Similarity	24.6%	Pred. No. 0.5;		
Matches 51, Conservative	27;	Mismatches	63;	Gaps 12.

RESULT 11  
 US-09-833-245-1567  
 Sequence 1567, Application US/09833245  
 Publication No. US20040010134A1  
 GENERAL INFORMATION:  
 APPLICANT: Human Genome Sciences, Inc.  
 TITLE OR INVENTION: Albumin Fusion Proteins  
 PRIORITY REFERENCE: PFS46PCT  
 CURRENT APPLICATION NUMBER: US/09/833,245  
 FILING DATE: 2001-04-12  
 CURRENT APPLICATION NUMBER: 60/229, 358

```

? PRIOR FILING DATE: 2000-04-12
? PRIOR APPLICATION NUMBER: 60/256, 931
? PRIOR FILING DATE: 2000-12-21
? PRIOR APPLICATION NUMBER: 60/199, 384
? PRIOR FILING DATE: 2000-04-25
? NUMBER OF SEQ ID NOS: 2267
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1567
? LENGTH: 255
? TYPE: PRT
? ORGANISM: Homo sapiens
? OS-09-813-245-1567

```

Query Match	6.5%;	Score 99;	DB 11;	Length 255;
Best Local Similarity	24.2%;	Pred. No. 0.18;		
Matches	40;	Conservative	22;	Mismatches 65;
				Indels 38;
				Gaps 8

```
QY      166 DSAILVEKSHPLAF-----CDT--LDNNEIVATIKF 197
          |   :   |   :   ||   :   |   :
Db      132 DVEYKFPQANSHFYFKYAQEDSGTYRCDVQLVKNLRLVKRLYF 176
          :
```

RESULT 12  
US-10-052-586-538

1	PRIOR FILING DATE: 1998-05-07
2	PRIOR APPLICATION NUMBER: 60/084640
3	PRIOR FILING DATE: 1998-05-07
4	PRIOR APPLICATION NUMBER: 60/084643
5	PRIOR FILING DATE: 1998-05-07
6	PRIOR APPLICATION NUMBER: 60/085573
7	PRIOR FILING DATE: 1998-05-15
8	PRIOR APPLICATION NUMBER: 60/085579
9	PRIOR FILING DATE: 1998-05-15
10	PRIOR APPLICATION NUMBER: 60/085580
11	PRIOR FILING DATE: 1998-05-15
12	PRIOR APPLICATION NUMBER: 60/085582
13	PRIOR FILING DATE: 1998-05-15
14	PRIOR APPLICATION NUMBER: 60/085700
15	PRIOR FILING DATE: 1998-05-15
16	PRIOR APPLICATION NUMBER: 60/086023
17	PRIOR FILING DATE: 1998-05-18
18	PRIOR APPLICATION NUMBER: 60/086392
19	PRIOR FILING DATE: 1998-05-22
20	PRIOR APPLICATION NUMBER: 60/086486
21	PRIOR FILING DATE: 1998-05-22
22	PRIOR APPLICATION NUMBER: 60/087098
23	PRIOR FILING DATE: 1998-05-28
24	PRIOR APPLICATION NUMBER: 60/087208
25	PRIOR FILING DATE: 1998-05-28
26	PRIOR APPLICATION NUMBER: 60/087609
27	PRIOR FILING DATE: 1998-06-02
28	PRIOR APPLICATION NUMBER: 60/087759
29	PRIOR FILING DATE: 1998-06-02
30	PRIOR APPLICATION NUMBER: 60/087827
31	PRIOR FILING DATE: 1998-06-03
32	PRIOR APPLICATION NUMBER: 60/088025
33	PRIOR FILING DATE: 1998-06-04
34	PRIOR APPLICATION NUMBER: 60/088028
35	PRIOR FILING DATE: 1998-06-04
36	PRIOR APPLICATION NUMBER: 60/088029
37	PRIOR FILING DATE: 1998-06-04
38	PRIOR APPLICATION NUMBER: 60/088033
39	PRIOR FILING DATE: 1998-06-04
40	PRIOR APPLICATION NUMBER: 60/088167
41	PRIOR FILING DATE: 1998-06-05
42	PRIOR APPLICATION NUMBER: 60/088202
43	PRIOR FILING DATE: 1998-06-05
44	PRIOR APPLICATION NUMBER: 60/088212
45	PRIOR FILING DATE: 1998-06-05
46	PRIOR APPLICATION NUMBER: 60/088217
47	PRIOR FILING DATE: 1998-06-05
48	PRIOR APPLICATION NUMBER: 60/088326
49	PRIOR FILING DATE: 1998-06-04
50	PRIOR APPLICATION NUMBER: 60/088655
51	PRIOR FILING DATE: 1998-06-09
52	PRIOR APPLICATION NUMBER: 60/088722
53	PRIOR FILING DATE: 1998-06-10
54	PRIOR APPLICATION NUMBER: 60/088738
55	PRIOR FILING DATE: 1998-06-10
56	PRIOR APPLICATION NUMBER: 60/088740
57	PRIOR FILING DATE: 1998-06-10
58	PRIOR APPLICATION NUMBER: 60/088811
59	PRIOR FILING DATE: 1998-06-10
60	PRIOR APPLICATION NUMBER: 60/088824
61	PRIOR FILING DATE: 1998-06-10
62	PRIOR APPLICATION NUMBER: 60/088825
63	PRIOR FILING DATE: 1998-06-10
64	PRIOR APPLICATION NUMBER: 60/088826
65	PRIOR FILING DATE: 1998-06-10
66	PRIOR APPLICATION NUMBER: 60/088861
67	PRIOR FILING DATE: 1998-06-11
68	PRIOR APPLICATION NUMBER: 60/088863
69	PRIOR FILING DATE: 1998-06-11
70	PRIOR APPLICATION NUMBER: 60/088876
71	PRIOR FILING DATE: 1998-06-11
72	PRIOR APPLICATION NUMBER: 60/089090
73	PRIOR FILING DATE: 1998-06-12

PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089908

Query Match 6.5%; Score 99; DB 13; Length 255;  
Best Local Similarity 24.2%; Pred. No. 0.18; Indels 38; Gaps 8;  
Matches 40; Conservative 22; Mismatches 65; Indels 38; Gaps 8;

QY 58 TAENYAPPEDEVSNVKEVEFGMCVTGIGVR-EVILNCGPGESKCVARVEBCR 116  
DB 25 TPKTLAIPKLOBAVAKVI--INATTCVTGCGYKEVEVCEVGPVGVRKCKOTQRLBCL 82  
QY 117 GPTDCGM-----GKPISLESVRLACHTSP-----NREKYMKLLRQDOOSIILVN 165  
DB 83 TWMICGMHLFTLLIGK-----EFELSCSSDILFEFGAEPFTWRLAR-----GVISTD 131  
QY 166 DSAILEVRKESHPLAF-----CDT--LDNNEIVATIKF 197  
DB 132 DEVFKPQANSHFVKRYAEOYDSGTYRCDVQLVKNLRVLYKLYF 176

## RESULT 13

US-10-174-590-538  
Sequence 538, Application US/10174590  
Publication No. US20030008352A1

GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C42  
CURRENT APPLICATION NUMBER: US/10/174,590  
CURRENT FILING DATE: 2002-06-18  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 538  
LENGTH: 255  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-174-590-538

Query Match 6.5%; Score 99; DB 14; Length 255;  
Best Local Similarity 24.2%; Pred. No. 0.18; Indels 38; Gaps 8;  
Matches 40; Conservative 22; Mismatches 65; Indels 38; Gaps 8;

QY 58 TAENYAPPEDEVSNVKEVEFGMCVTGIGVR-EVILNCGPGESKCVARVEBCR 116  
DB 25 TPKTLAIPKLOBAVAKVI--INATTCVTGCGYKEVEVCEVGPVGVRKCKOTQRLBCL 82  
QY 117 GPTDCGM-----GKPISLESVRLACHTSP-----NREKYMKLLRQDOOSIILVN 165  
DB 83 TWMICGMHLFTLLIGK-----EFELSCSSDILFEFGAEPFTWRLAR-----GVISTD 131  
QY 166 DSAILEVRKESHPLAF-----CDT--LDNNEIVATIKF 197

DB 132 DEVFKPQANSHFVKRYAEOYDSGTYRCDVQLVKNLRVLYKLYF 176

## RESULT 14

US-10-176-758-538  
Sequence 538, Application US/10176758  
Publication No. US20030008353A1

GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C104  
CURRENT APPLICATION NUMBER: US/10/176,758  
CURRENT FILING DATE: 2002-06-21  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 538  
LENGTH: 255  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-758-538

Query Match 6.5%; Score 99; DB 14; Length 255;  
Best Local Similarity 24.2%; Pred. No. 0.18; Indels 38; Gaps 8;  
Matches 40; Conservative 22; Mismatches 65; Indels 38; Gaps 8;

QY 58 TAENYAPPEDEVSNVKEVEFGMCVTGIGVR-EVILNCGPGESKCVARVEBCR 116  
DB 25 TPKTLAIPKLOBAVAKVI--INATTCVTGCGYKEVEVCEVGPVGVRKCKOTQRLBCL 82  
QY 117 GPTDCGM-----GKPISLESVRLACHTSP-----NREKYMKLLRQDOOSIILVN 165  
DB 83 TWMICGMHLFTLLIGK-----EFELSCSSDILFEFGAEPFTWRLAR-----GVISTD 131  
QY 166 DSAILEVRKESHPLAF-----CDT--LDNNEIVATIKF 197  
DB 132 DEVFKPQANSHFVKRYAEOYDSGTYRCDVQLVKNLRVLYKLYF 176

## RESULT 15

US-10-175-737-538  
Sequence 538, Application US/10175737  
Publication No. US20030013153A1

GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C50  
CURRENT APPLICATION NUMBER: US/10/175,737  
CURRENT FILING DATE: 2002-06-19  
Prior application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 612

```

; SEQ ID NO 538
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-538

```

Query Match	6.5%	Score 99	DB 14	Length 255
Best Local Similarity	24.2%	Pred. No. 0.18		
Matches 40	Conservative 22	Mismatches 65	Indels 38	Gaps 8

```

QY  TAENYAPETEDVSRRNYKVEYFPGMTCYTCGAGV--EYLINNGCPGSGSKCVARVEEOR  116
Db  25  TKPTLAIPEKLOBAVGKVI--INATTCYTCGAGYKEETVCEYGPBGVRRKCTQALBCL  82
QY  117  GPTDCGW-----GKPISESLSEYALACIHTSPV---NRFKYMKLRLRODOSIILVN  165
Db  83  TNNWICGMHLFTLLIG-----EPBLSTCLSDILEFQGEARFRTYRLAR-----GVISTD  131
QY  166  DSAILEVRKESHPLAF-----CDT--LDNNEIVATIKF  197
Db  132  DEVKPEPQANSHFVAFKRYAQVEYDSGYRRCDVGLVKRLAIRVKRLYF  176

```

Search completed: April 15, 2005, 14:18:51  
Job time : 138 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2005, 14:04:21; Search time 173 Seconds  
(without alignment)  
657.269 Million cell updates/sec

Title: US-10-809-655-9

Perfect score: 1528

Sequence: 1 MSPRGTGCSAGLMTVGVLT.....LDQLPTMPSGDALSEWNE 294

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: A. Geneseq\_16Dec04:\*

1: geneseqp19808:\*\n2: geneseqp19908:\*\n3: geneseqp20008:\*\n4: geneseqp20018:\*\n5: geneseqp20028:\*\n6: geneseqp20038:\*\n7: geneseqp20038:\*\n8: geneseqp20048:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1528	100.0	294	AA85346 Sperm spe
2	762.5	49.9	163	ABP62052 Human sec
3	762.5	49.9	163	ADA57229 Human sec
4	762.5	49.9	163	ADA41111 Human sec
5	762.5	49.9	163	ADA37978 Human sec
6	593	38.8	124	ABP62066 Human sec
7	593	38.8	124	ADA57562 Human sec
8	593	38.8	124	ADA41456 Human sec
9	593	38.8	124	ADA38087 Human sec
10	593	38.8	125	AA12953 Amino aci
11	110	7.2	1315	ADN20693 Bacteriol
12	103	6.7	255	AA31833 Human foe
13	103	6.7	255	ABG02357 Novel hum
14	103	6.7	255	AAU96170 Human sec
15	103	6.7	255	ABG64819 Human alb
16	103	6.7	255	ADL78086 Albumin f
17	102	6.7	1083	ABR61710 Drosophili
18	99	6.5	255	AA873684 Human oxi
19	99	6.5	255	AAU29232 Human PRO
20	99	6.5	255	AAU96195 Human sec
21	99	6.5	255	ABG64818 Human alb
22	99	6.5	255	ABU58668 Human PRO
23	99	6.5	255	ABU88216 Novel hum
24	99	6.5	255	ABU84531 Human sec
25	59	6.5	255	ABR6405 Human sec

26	99	6.5	255	ABR65795	Abt65795 Human sec
27	99	6.5	255	ABU99735	Abu99735 Human sec
28	99	6.5	255	ABU82974	Abu82974 Human PRO
29	99	6.5	255	ABU90095	Abu90095 Novel hum
30	99	6.5	255	ABR68344	Abt68344 Human sec
31	99	6.5	255	ABU96397	Abu96397 Novel hum
32	99	6.5	255	ABU92828	Abu92828 Human sec
33	99	6.5	255	ABO08905	ABO08905 Human sec
34	99	6.5	255	ABO02957	ABO02957 Human sec
35	99	6.5	255	ABR75111	Abt75111 Human sec
36	99	6.5	255	ABR94873	Abt94873 Human sec
37	99	6.5	255	ABU85846	Abu85846 Human PRO
38	99	6.5	255	ABU99006	Abu99006 Novel hum
39	99	6.5	255	ABU98221	Abu98221 Novel hum
40	99	6.5	255	ABU91927	Abu91927 Novel hum
41	99	6.5	255	ABU89620	Abu89620 Human PRO
42	99	6.5	255	ABU86461	Abu86461 Human sec
43	99	6.5	255	ABU67674	Abu67674 Human sec
44	99	6.5	255	ABU80702	Abu80702 Human PRO
45	99	6.5	255	ABR99620	Abt99620 Human sec

## ALIGNMENTS

RESULT 1  
ID AAB85346 standard; protein, 294 AA.  
XX AAB85346;  
AC XX  
DT 17-SEP-2001 (first entry)  
DT XX  
DE Sperm specific surface protein SMP32.  
XX Sperm specific surface protein; C7/8; SMP32; C58; contraceptive;  
XX vaccine; antifertility; spermicide.  
XX  
OS Homo sapiens.  
XX WO200153352-A2.  
XX  
XX 26-JUL-2001.  
XX  
XX 19-JAN-2001; 2001WO-US001717.  
XX  
XX 19-JAN-2000; 2000US-0176885P.  
XX  
XX (UUYI-) UNIV VIRGINIA PATENT FOUND.  
XX  
XX Hao Z, Herr JC, Jayes FL, Shetty J, Wolkowicz MJ;  
XX WPI; 2001-451902/48.  
XX N-PSDB; AAH22947.  
XX  
XX New human sperm surface proteins C7/8, SMP32 and C58 for development of  
XX contraceptive vaccines.  
XX  
XX Claim 10; Page 56-67; 63pp; English.  
XX  
XX The invention relates to novel human sperm specific surface proteins,  
XX C7/8, SMP32 and C58. The proteins, nucleic acids encoding the sperm  
XX specific surface proteins and antibodies specific for the proteins are  
XX useful for making contraceptive compositions including contraceptive  
XX vaccines. The vaccines produced by the invention are hoped to be more  
XX effective than previous contraceptive vaccines which only caused 75%  
XX inhibition of fertility. The present sequence represents the human sperm  
XX specific surface protein SMP32  
XX  
XX Sequence 294 AA;  
XX  
XX Query Match 100.0%; Score 1528; DB 4; Length 294;  
XX Best Local Similarity 100.0%; Pred. No. 3.2e-157;

Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPRGTCGSAGLMTVMGMLLAGIQAARGNTVTAQVADAGLAHGEGBEETENDSETRA 60  
 DB 1 MSPRGTCGSAGLMTVMGMLLAGIQAARGNTVTAQVADAGLAHGEGBEETENDSETRA 60  
 QY 61 NYAPPEDEVNRRNVKVEFGMCTVTCGIGVREVIILNCGPGGSKCVAVBEBCRGPTD 120  
 DB 61 NYAPPEDEVNRRNVKVEFGMCTVTCGIGVREVIILNCGPGGSKCVAVBEBCRGPTD 120  
 QY 121 CGWGPISLSLSVRLACIHTSPINRFKYMKMLRQDOOSIILVNDAILVEKESHPILA 180  
 DB 121 CGWGPISLSLSVRLACIHTSPINRFKYMKMLRQDOOSIILVNDAILVEKESHPILA 180  
 QY 181 PECTLDNNEIVATIKFTVYTSSELQMRSSLPATDAILPVLITGVIIICVPIIFLLFI 240  
 DB 181 PECTLDNNEIVATIKFTVYTSSELQMRSSLPATDAILPVLITGVIIICVPIIFLLFI 240  
 QY 241 IINNAVAVAFGAKASTPEVQSEGSVRYKOSTSIDOLPTMPGEBDMLSEWNE 294  
 DB 241 IINNAVAVAFGAKASTPEVQSEGSVRYKOSTSIDOLPTMPGEBDMLSEWNE 294

## RESULT 2

ABP62052  
 ID ABP62052 standard; protein; 163 AA.

AC ABP62052;

DT 12-NOV-2002 (first entry)

DE Human secreted protein SEQ ID NO 105.

XX Human; noctropic; neuroprotective; cyrostatic; dermatological; virucide;  
 XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
 XX antiparkinsonian; antislaking; antianaemic; antiallergic; cancer;  
 XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
 XX antiallergic; antidiabetic; antidiacer; anticonvulsant; antifungal;  
 XX antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
 XX neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

XX OS WO200257420-A2.

XX PD 25-JUL-2002.

XX 17-JAN-2002; 2002WO-US001109.

XX PR 18-JAN-2001; 2001US-0262066P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen CA, Olsen H;  
 XX Ebner R, Brewer LA;

XX DR MPI; 2002-599716/64.

XX N-PSDB; ABQ92592.

XX New polynucleotides and polypeptides useful for diagnosing, prognosing,  
 XX treating or preventing e.g. neurodegenerative, central nervous system,  
 XX autoimmune, respiratory, reproductive, or inflammatory diseases or  
 XX disorders.

XX Claim 11, Page 749-750; 785pp; English.

XX The invention relates to novel genes (ABQ92553-ABQ92607) and proteins  
 XX (ABP62013-ABP62153) useful for preventing, treating or ameliorating  
 XX medical conditions e.g. by protein or gene therapy. The genes are  
 XX isolated from a range of human tissues disclosed in the specification.  
 XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,

CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
 CC / (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections  
 XX  
 SQ Sequence 163 AA;

Query Match 49.9%; Score 762.5; DB 5; Length 163;  
 Best Local Similarity 94.2%; Pred.No. 3.2e-74;  
 Matches 147; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MSPRGTCGSAGLMTVMGMLLAGIQAARGNTVTAQVADAGLAHGEGBEETENDSETRA 60  
 DB 1 MSPRGTCGSAGLMTVMGMLLAGIQAARGNTVTAQVADAGLAHGEGBEETENDSETRA 60  
 QY 61 NYAPPEDEVNRRNVKVEFGMCTVTCGIGVREVIILNCGPGGSKCVAVBEBCRGPTD 120  
 DB 61 NYAPPEDEVNRRNVKVEFGMCTVTCGIGVREVIILNCGPGGSKCVAVBEBCRGPTD 120  
 QY 121 CGWGPISLSLSVRLACIHTSPINRFKYMKMLRQ 156  
 DB 121 CGWGPISLSLSVRLACIHTSPINRFKYMKMLRQ 155

## RESULT 3

ADA57229  
 ID ADA57229 standard; protein; 163 AA.

AC ADA57229;

DT 20-NOV-2003 (first entry)

DE Human secreted protein #512.

XX immunosuppressive; antiinflammatory; antiaesthetic; antiallergic;  
 XX cyrostatic; cerebroprotective; neuroprotective; noctropic;  
 XX cardiovascular; antiatherosclerotic; gene therapy;  
 XX human secreted protein; immune disorder; inflammation;  
 XX respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;  
 XX inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;  
 XX multiple sclerosis; ischaemic brain injury; Parkinson's disease;  
 XX Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;  
 XX triple helix formation; antisense gene therapy; forensic biology.

XX Homo sapiens.

XX OS WO2002102994-A2.

XX PD 27-DEC-2002.

XX 19-MAR-2002; 2002WO-US008278.

XX PR 21-MAR-2001; 2001US-0277340P.

XX PR 19-JUL-2001; 2001US-0306171P.

XX PR 13-NOV-2001; 2001US-0331287P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR MPI; 2003-167512/16.

XX N-PSDB; ADA56333.

XX New human secreted polypeptides and polynucleotides, useful for  
 XX diagnosing, treating or preventing e.g. immune disorders, inflammatory  
 XX conditions, respiratory disorders, cancers, CNS disorders, or  
 XX neurodegenerative disorders.

XX Claim 13, SEQ ID NO 1419; 1754pp; English.

XX The invention relates to 592 new human secreted polypeptides useful for  
 CC diagnosing, treating or preventing e.g. immune disorders, inflammatory  
 CC conditions, respiratory disorders, cancers, CNS disorders, or  
 CC neurodegenerative disorders, or polypeptides comprising an amino acid  
 CC sequence at least 95% identical to the new sequences. The polypeptides,  
 CC antibodies or antibody fragments that bind to the polypeptides, nucleic  
 CC acids encoding the polypeptides, agonists or antagonists that binds to  
 CC the polypeptide, are useful in preparing diagnostic or pharmaceutical  
 CC compositions for diagnosing, treating or preventing an e.g. immune  
 CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,  
 CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and  
 CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders  
 CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative  
 CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and  
 CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The  
 CC polynucleotides are useful for chromosome identification, chromosome  
 CC mapping, for controlling gene expression through triple helix formation  
 CC or antisense DNA or RNA, in gene therapy, for identifying individuals  
 CC from minute biological samples, in forensic biology, and as hybridization  
 CC probes. The polypeptides are useful for as molecular weight markers on  
 CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)  
 CC gels, to raise antibodies, for testing biological activities, and for  
 CC treating or preventing neural disorders, immune system disorders,  
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
 CC renal, proliferative and/or cancerous diseases. This sequence corresponds  
 CC to one of the polypeptide of the invention. Note: The sequence data for  
 CC this patent did form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 163 AA:

XX Query Match 49.9%; Score 762.5; DB 6; Length 163;  
 XX Best Local Similarity 94.2%; Pred. No. 3.2e-74;  
 XX Matches 147; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MSPRGTGSAGILMTVWGLLAGLQSGARGNTVAOVQDAGLAHBEGBEETNNSETAE 60  
 DB 1 MSPRGTGSAGILMTVWGLLAGLQSGARGNTVAOVQDAGLAHBEGBEETNNSETAE 60  
 QY 61 NYAPETEDVSNRNVKEVEFGMCVTGIGVREVLITNGCGSGSKCVARVEBCRGPTD 120  
 DB 61 NYAPETEDVSNRNVKEVEFGMCVTGIGVREVLITNGCGSGSKCVARVEBCRGPTD 120  
 QY 121 CGMGKPISESLSVRLACIHTSPILNRFKYMKLRLQ 156  
 DB 121 CGMGKPISESLSVRLACIHTSPILNRFKYMKLRLQ 156  
 DB 121 CGMGKPISESLSVRLACIHTSPILNRFKYMKLRLQ 155

XX RESULT 4  
 XX ADA41111  
 XX ID ADA41111 standard; protein; 163 AA.

XX ADA41111;  
 XX DT 20-NOV-2003 (first entry)

XX Human secreted protein.  
 XX Human, secreted protein; cancer; hyperproliferative disorder;  
 XX Rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;  
 XX anaemia; allergic reaction; asthma; cardiovascular disorder;  
 XX wound healing; cytotoxic; immunosuppressive; neuroprotective;  
 XX antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;  
 XX vulnerability; candiant; gene therapy.

XX Homo sapiens.  
 XX AD37978  
 XX ID W02002102393-A2.  
 XX AC 27-DEC-2002.

PF 19-MAR-2002; 2002MO-US008123.  
 XX 21-MAR-2001; 2001US-0277340P.  
 PR 19-JUL-2001; 2001US-0306171P.  
 PR 13-NOV-2001; 2001US-0331287P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Rosen CA, Ruben SM;  
 XX WPI; 2003-175238/17.  
 DR New human secreted proteins and nucleic acid molecules, useful for  
 XX preparing a diagnostic or pharmaceutical composition for diagnosing,  
 XX preventing or treating cancer or other hyperproliferative disorder,  
 XX asthma, allergies or AIDS.

PS Claim 1; SEQ ID NO 1493; 3205PP; English.

XX The invention relates to novel genes ADA39629-ADA40565 and proteins  
 CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,  
 CC treating or ameliorating medical conditions e.g. by protein or gene  
 CC therapy. The polypeptides, nucleic acid molecules, antibodies or their  
 CC fragments, and agonists or antagonists that bind to the polypeptide are  
 CC useful for preparing a diagnostic or pharmaceutical composition for  
 CC diagnosing or treating cancer or other hyperproliferative disorder. The  
 CC polypeptides and nucleic acid molecules are also useful for detecting,  
 CC preventing, diagnosing, prognosticating, treating or ameliorating cancer  
 CC or other hyperproliferative disorders including neoplasms, autoimmune  
 CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus  
 CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic  
 CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,  
 CC thrombocytopenia), allergic reactions including asthma or eczema,  
 CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory  
 CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.  
 CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders  
 CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,  
 CC fungal or viral infections including HIV/AIDS), or wound healing and  
 CC disorders of epithelial cell proliferation. The nucleic acids are also  
 CC useful for chromosome identification, radiation hybrid mapping or long-  
 CC range restriction mapping, as molecular weight markers, or as  
 CC hybridization or diagnostic probes. The polypeptides and antibodies are  
 CC useful for providing immunological probes for differential identification  
 CC of the tissues immunohistochemistry assays. Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 163 AA:

XX Query Match 49.9%; Score 762.5; DB 6; Length 163;  
 XX Best Local Similarity 94.2%; Pred. No. 3.2e-74;  
 XX Matches 147; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MSPRGTGSAGILMTVWGLLAGLQSGARGNTVAOVQDAGLAHBEGBEETNNSETAE 60  
 DB 1 MSPRGTGSAGILMTVWGLLAGLQSGARGNTVAOVQDAGLAHBEGBEETNNSETAE 60  
 QY 61 NYAPETEDVSNRNVKEVEFGMCVTGIGVREVLITNGCGSGSKCVARVEBCRGPTD 120  
 DB 61 NYAPETEDVSNRNVKEVEFGMCVTGIGVREVLITNGCGSGSKCVARVEBCRGPTD 120  
 QY 121 CGMGKPISESLSVRLACIHTSPILNRFKYMKLRLQ 156  
 DB 121 CGMGKPISESLSVRLACIHTSPILNRFKYMKLRLQ 155

XX RESULT 5  
 XX ADD37978  
 XX ID ADD37978 standard; protein; 163 AA.

XX ADD37978;  
 XX AC

DT 15-JAN-2004 (first entry)  
XX Human secreted protein #161.  
DB human secreted protein; Antiallergic; Antiinflammatory; Antibacterial;  
KW Anti-HIV; Cytostatic; Immunosuppressive; Hemostatic.  
XX Homo sapiens.  
OS  
XX MO200290526-A2.  
PN  
XX 14-NOV-2002.  
PD  
XX 19-MAR-2002; 2002WO-US008279.  
PF  
XX 21-MAR-2001; 2001US-0277340P.  
PR 19-JUL-2001; 2001US-0306171P.  
PR 13-NOV-2001; 2001US-0331287P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Ruben SM;  
PI  
XX WPI; 2003-140218/13.  
DR  
XX New human secreted proteins and nucleic acid molecules, useful for  
PT preparing a diagnostic or pharmaceutical composition for diagnosing or  
PT treating allergic or asthmatic disorders, or related immediate  
PT hypersensitivity disorders.  
XX  
XX Claim 1; SEQ ID NO 460; 1323pp; English.  
PS  
XX The present invention relates to an isolated polypeptide or human  
CC secreted protein. The polypeptides, nucleic acid molecules, antibodies or  
CC their fragments, and agonists or antagonists that bind are useful for  
CC preparing a diagnostic or pharmaceutical composition for diagnosing or  
CC treating allergic or asthmatic disorders. The polypeptide is also useful  
CC for identifying a binding partner by contacting the polypeptide with a  
CC binding partner, and determining whether the binding partner increases or  
CC decreases the activity of the polypeptide. The polypeptides and nucleic  
CC acid molecules are also useful for detecting, preventing, diagnosing,  
CC prognosticating, treating or ameliorating inflammatory disorders  
CC neoplastic diseases, wound healing and disorders of epithelial cell  
CC proliferation, immune disorders, cardiovascular disorders, blood-related  
CC disorders, infectious diseases, endocrine disorders, or gastrointestinal  
CC disorders. The nucleic acids are also useful for chromosome  
CC identification, radiation hybrid mapping or long-range restriction  
CC mapping, as molecular weight markers, or as hybridization or diagnostic  
CC probes. The polypeptides and antibodies are useful for providing  
CC immunological probes for differential identification of the tissues  
CC immunohistochemistry assays. The present sequence represents a human  
CC secreted protein.  
XX  
XX Sequence 163 AA;  
SQ  
Query Match 49.9%; Score 762.5; DB 7; Length 163;  
Best Local Similarity 94.2%; Pred. No. 3.2e-74;  
Matches 147; Conservative 2; Mismatches 6; Indels 1; Gaps 1;  
QY 1 MSPRGTGSAAGLMTVGMILLAGLQASAGTNTAAVQDAGLAHGBEGEETENNDSYAE 60  
DB 1 MSPRGTGSAAGLMTVGMILLAGLQASAGTNTAAVQDAGLAHGBEGEETENNDSYAE 60  
QY 61 NYAPETEDVSNRNVAKEVEFGMCTVTCGIGREAVILITNGCGPGBSKCVAVVEECRPTD 120  
DB 61 NYAPETEDVSNRNVAKEVEFGMCTVTCGIGREAVILITNGCGPGBSKCVAVVEECRPTD 120  
QY 121 CGMKPISSESVRLACHTSPILARFKMMMLLQ 156  
DB 121 CGMKPISSESVRLACHTSPILVSTV-ELLQ 155

RESULT 6

ABP62066  
ID ABP62066 standard; protein; 124 AA.  
XX  
XX AC ABP62066;  
XX  
XX 12-NOV-2002 (first entry)  
XX  
XX Human secreted protein SEQ ID NO 119.  
DE  
XX Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;  
KW immunosuppressive; anti-infective; anti-HIV; antibacterial; vulnery;  
KW antiparkinsonian; antitoxic; antianaemic; antiarthritic; cancer;  
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.  
XX Homo sapiens.  
OS  
XX MO200257420-A2.  
PN  
XX 25-JUL-2002.  
PD  
XX 17-JAN-2002; 2002WO-US001109.  
PF  
XX 18-JAN-2001; 2001US-0262066P.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Moore PA, Ruben SM, Lafleur DW, Shi Y, Rosen CA, Olsen H;  
PI Ebner R, Brewer LA;  
XX WPI; 2002-599716/64.  
DR  
XX New polynucleotides and polypeptides useful for diagnosing, prognosing,  
PT treating or preventing e.g. neurodegenerative, central nervous system,  
PT autoimmune, respiratory, reproductive, or inflammatory diseases or  
PT disorders.  
XX  
XX Claim 11; Page 754-755; 785pp; English.  
PS  
XX The invention relates to novel genes (ABQ92553-ABQ92607) and proteins  
CC (ABP62013-ABP62153) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections  
XX  
XX Sequence 124 AA;  
SQ  
Query Match 38.8%; Score 593; DB 5; Length 124;  
Best Local Similarity 98.3%; Pred. No. 6.1e-56;  
Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MSPRGTGSAAGLMTVGMILLAGLQASAGTNTAAVQDAGLAHGBEGEETENNDSYAE 60  
DB 1 MSPRGTGSAAGLMTVGMILLAGLQASAGTNTAAVQDAGLAHGBEGEETENNDSYAE 60  
QY 61 NYAPETEDVSNRNVAKEVEFGMCTVTCGIGREAVILITNGCGPGBSKCVAVVEEC 115  
DB 61 NYAPETEDVSNRNVAKEVEFGMCTVTCGIGREAVILITNGCGPGBSKCVAVVEEC 115

RESULT 7

AD57562  
ID ADA57562 standard; protein; 124 AA.  
XX  
XX ADA57562;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX Human secreted protein #512.  
XX  
XX immunosuppressive; antiinflammatory; antiaesthetic; antiallergic;  
XX cytostatic; cerebroprotective; neuroprotective; nootropic;  
XX cardiovascular; antiarteriosclerotic; gene therapy;  
XX human secreted protein; immune disorder; inflammation;  
XX respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;  
XX inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;  
XX multiple sclerosis; ischaemic brain injury; Parkinson's disease;  
XX Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;  
XX triplex helix formation; antisense gene therapy; forensic biology.  
XX  
XX Homo sapiens.  
XX  
XX WO2002102994-A2.  
XX  
XX 27-DEC-2002.  
XX  
XX 19-MAR-2002; 2002WO-US008278.  
XX  
XX 21-MAR-2001; 2001US-0277340P.  
XX 19-JUL-2001; 2001US-0306171P.  
XX 13-NOV-2001; 2001US-0331287P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2003-167512/16.  
XX N-PSDB; ADA56669.  
XX  
XX New human secreted polypeptides and polynucleotides, useful for  
XX diagnosing, treating or preventing e.g. immune disorders, inflammatory  
XX conditions, respiratory disorders, cancers, CNS disorders, or  
XX neurodegenerative disorders, or polypeptides comprising an amino acid  
XX sequence at least 95% identical to the new sequences. The polypeptides,  
XX antibodies or antibody fragments that bind to the polypeptides, nucleic  
XX acids encoding the polypeptides, agonists or antagonists that binds to  
XX the polypeptide, are useful in preparing diagnostic or pharmaceutical  
XX compositions for diagnosing, treating or preventing an e.g. immune  
XX disorders, inflammatory conditions (e.g. inflammatory bowel disease,  
XX nephritis or Crohn's disease), respiratory disorders (e.g. asthma and  
XX allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders  
XX (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative  
XX disorders (e.g. Parkinson's disease or Alzheimer's disease), and  
XX cardiovascular disorders (e.g. atherosclerosis or myocarditis). The  
XX polynucleotides are useful for chromosome identification, chromosome  
XX mapping, for controlling gene expression through triple helix formation  
XX or antisense DNA or RNA, in gene therapy, for identifying individuals  
XX from minute biological samples, in forensic biology, and as hybridization  
XX probes. The polypeptides are useful for as molecular weight markers on  
XX sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)  
XX gels, to raise antibodies, for testing biological activities, and for  
XX treating or preventing neural disorders, immune system disorders,  
XX muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
XX renal, proliferative and/or cancerous diseases. This sequence corresponds  
XX to one of the polypeptide of the invention. Note: The sequence data for  
XX this patent did form part of the printed specification, but was obtained  
XX in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 124 AA;  
SQ  
Query Match 38.8%; Score 593; DB 6; Length 124;  
Best Local Similarity 98.3%; Pred. No. 6,1e-56;  
Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MSPRGTCGAGLMTVMGILLAGLSARGTNTAIVADAGLHBECEETENNSETAE 60  
DB 1 MSPRGTCGAGLMTVMGILLAGLSARGTNTAIVADAGLHBECEETENNSETAE 60  
QY 61 NYAPETEDVSNRNKVKVEFGMCTVTCGIGREYILTNCGCGSSKCVRVEEC 115  
DB 61 NYAPETEDVSNRNKVKVEFGMCTVTCGIGREYILTNCGCGSSKCVRVEEC 115  
RESULT 8  
ADA1456  
ID ADA1456 standard; protein; 124 AA.  
XX  
XX ADA1456;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX Human secreted protein.  
XX  
XX Human; secreted protein; cancer; hyperproliferative disorder;  
XX Rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;  
XX anaemia; allergic reaction; asthma; cardiovascular disorder;  
XX wound healing; cytostatic; immunosuppressive; nootropic; neuroprotective;  
XX antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;  
XX vulnery; cardiant; gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX WO2002102993-A2.  
XX  
XX 27-DEC-2002.  
XX  
XX 19-MAR-2002; 2002WO-US008123.  
XX  
XX 21-MAR-2001; 2001US-0277340P.  
XX 19-JUL-2001; 2001US-0306171P.  
XX 13-NOV-2001; 2001US-0331287P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2003-175238/17.  
XX  
XX New human secreted proteins and nucleic acid molecules, useful for  
XX preparing a diagnostic or pharmaceutical composition for diagnosing,  
XX preventing or treating cancer or other hyperproliferative disorder,  
XX asthma, allergies or AIDS.  
XX  
XX Claim 1; SEQ ID NO 1839; 3205pp; English.  
XX  
XX The invention relates to novel genes ADA39629-ADA40565 and proteins  
XX ADA40566-ADA41501 for human secreted proteins, useful for preventing,  
XX treating or ameliorating medical conditions e.g. by protein or gene  
XX therapy. The polypeptides, nucleic acid molecules, antibodies or their  
XX fragments, and agonists or antagonists that bind to the polypeptide are  
XX useful for preparing a diagnostic or pharmaceutical composition for  
XX diagnosing or treating cancer or other hyperproliferative disorder. The  
XX polypeptides and nucleic acid molecules are also useful for detecting,  
XX preventing, diagnosing, prognosticating, treating or ameliorating cancer  
XX or other hyperproliferative disorders including neoplasms, autoimmune  
XX disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus  
XX erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic  
XX anaemia), haematopoietic or haematological disorders (e.g. anaemia,  
XX thrombocytopenia), allergic reactions including asthma or eczema.

inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory bowel disease or Crohn's disease), neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's disease), cardiovascular disorders (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial, fungal or viral infections including HIV/AIDS), or wound healing and disorders of epithelial cell proliferation. The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping, as molecular weight markers, or as hybridization or diagnostic probes. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 124 AA;

Query Match 38.8%; Score 593; DB 6; Length 124;  
Best Local Similarity 98.3%; Pred. No. 6.1e-56;  
Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MSPRGTCGAGLMTVGMILLAGLQSGAGTNTAAVODAGLAHSGEGEETENDSETAE 60  
1 MSPRGTCGAGLMTVGMILLAGLQSGAGTNTAAVODAGLAHSGEGEETENDSETAE 60  
61 NYAPETEDVSNRNKXKVEFGMCTVTCGIGREVILLTNGCGSGSKCVAVVEEC 115  
61 NYAPETEDVSNRNKXKVEFGMCTVTCGIGREVILLTNGCGSGSKCVAVVEEC 115

RESULT 9

ADD38087  
ADD38087 standard; protein; 124 AA.

ADD38087;

15-JAN-2004 (first entry)

Human secreted protein #270.

Human secreted protein; Antiallergic; Antiinflammatory; Antibacterial;

Anti-HIV; Cytostatic; Immunosuppressive; Hemostatic.

Homo sapiens.

MO200290526-A2.

14-NOV-2002.

19-MAR-2002; 2002MO-US008279.

21-MAR-2001; 2001US-0277340P.

19-JUL-2001; 2001US-0306171P.

13-NOV-2001; 2001US-0331287P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruden SM,

WPI; 2003-140218/13.

New human secreted proteins and nucleic acid molecules, useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating allergic or asthmatic disorders, or related immediate hypersensitivity disorders.

Claim 1; SEQ ID NO 569; 1323bp; English.

The present invention relates to an isolated polypeptide or human secreted protein. The polypeptides, nucleic acid molecules, antibodies or their fragments, and agonists or antagonists that bind are useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating allergic or asthmatic disorders. The polypeptide is also useful

for identifying a binding partner by contacting the polypeptide with a binding partner, and determining whether the binding partner increases or decreases the activity of the polypeptide. The polypeptides and nucleic acid molecules are also useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating inflammatory disorders, neoplastic diseases, wound healing and disorders of epithelial cell proliferation, immune diseases, cardiovascular disorders, blood-related disorders, infectious diseases, endocrine disorders, or gastrointestinal disorders. The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping, as molecular weight markers, or as hybridization or diagnostic probes. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. The present sequence represents a human secreted protein.

Sequence 124 AA;

Query Match 38.8%; Score 593; DB 7; Length 124;  
Best Local Similarity 98.3%; Pred. No. 6.1e-56;  
Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MSPRGTCGAGLMTVGMILLAGLQSGAGTNTAAVODAGLAHSGEGEETENDSETAE 60  
1 MSPRGTCGAGLMTVGMILLAGLQSGAGTNTAAVODAGLAHSGEGEETENDSETAE 60  
61 NYAPETEDVSNRNKXKVEFGMCTVTCGIGREVILLTNGCGSGSKCVAVVEEC 115  
61 NYAPETEDVSNRNKXKVEFGMCTVTCGIGREVILLTNGCGSGSKCVAVVEEC 115

RESULT 10

AA12953  
AA12953 standard; protein; 125 AA.

AA12953;

17-JUN-1999 (first entry)

Amino acid sequence of a human secreted peptide.

Human secreted protein; cancer; immune disorder; infection;

inflammatory disorder; skin disorder; tumour; atherosclerosis;

restenosis; autoimmune disorder; Alzheimer's disease;

peripheral neuropathy; trauma; spinal cord injury; allergy;

hematopoietic disorder; skeletal disorder; neurological disorder;

atrial fibrillation; asthma; immunodeficiency disease; AIDS;

transplant rejection; ss.

Homo sapiens.

MO9911293-A1.

11-MAR-1999.

03-SEP-1998; 98MO-US018360.

05-SEP-1997; 97US-0057626P.

05-SEP-1997; 97US-0057632P.

05-SEP-1997; 97US-0057639P.

12-SEP-1997; 97US-0058666P.

12-SEP-1997; 97US-0058973P.

12-SEP-1997; 97US-0058974P.

(HUMA-) HUMAN GENOME SCI INC.



XX Human foetal kidney secreted protein m310\_1.  
 XX Secreted protein; m310\_1; human; therapy; diagnosis; vaccine; kidney.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX Domain 20  
 XX /note= "a putative transmembrane domain is centered  
 XX around this residue"  
 XX Domain 90  
 XX /note= "a putative transmembrane domain is centered  
 XX around this residue"  
 XX Domain 230  
 XX /note= "a putative transmembrane domain is centered  
 XX around this residue"  
 XX  
 XX MO9947555-A1.  
 XX  
 XX 23-SEP-1999.  
 XX  
 XX 18-MAR-1999; 99WO-US005939.  
 XX  
 XX 20-MAR-1998; 98US-0078603P.  
 XX 17-MAR-1999; 99US-00078803.  
 XX  
 XX (GENY ) GENETICS INST INC.  
 XX  
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 XX Merberg D, Treacy M, Agoetino MJ, Steinger RJ;  
 XX  
 XX WPI, 1999-562059/47.  
 XX DR N-PSDB; AA219897.  
 XX  
 XX New polynucleotides derived from murine fetal cell cDNA libraries,  
 XX potentially used as, e.g. vaccines.  
 XX  
 XX Claim 19(a); Page 98; 107pp; English.  
 XX  
 XX This is the predicted amino acid sequence of a novel human secreted  
 XX protein, m310\_1, as deduced from an isolated foetal kidney (293 cell)  
 XX clone (see AA219897). The invention provides new human secreted proteins  
 XX (see AA219828-38) and polynucleotides (see AA219893-901) isolated from  
 XX foetal cell, adult blood, adult brain and foetal kidney cDNA libraries.  
 XX They are predicted to have biological activities which would make them  
 XX suitable for treating, preventing or ameliorating medical conditions in  
 XX humans and animals, although no supporting data are given. Suggested  
 XX activities include nutritional, cytokine, tissue growth, cell  
 XX proliferation and differentiation, immunostimulant (e.g. as vaccine),  
 XX immunosuppressive, haematopoiesis regulating, activin or inhibin,  
 XX chemotactic or chemokinetic, haemostatic or thrombolytic, receptor/ligand  
 XX activity, antiinflammatory, cachexin or tumour invasion suppressor, and  
 XX tumour inhibition activities  
 XX  
 XX Sequence 255 AA;  
 XX  
 XX Query Match 6.7%; Score 103; DB 2; Length 255;  
 XX Best Local Similarity 24.8%; Pred. No. 0.042;  
 XX Matches 41; Conservative 21; Mismatches 65; Indels 38; Gaps 8;  
 XX  
 XX 58 TAENAPAPETEDVSRNRYKVEFGMCVTGIGVR-EVILNCGCGSGKCVVVEGR 116  
 XX 25 TPKTALPEKIOEAVGKVI--INATTCVTGIGLKEETVCEVGPDRKCKQTRLECL 82  
 XX 117 GPTDCG-----GKPISESLSEVRLACIHTSP---NRFKYMVKLLRDOOSIILVN 165  
 XX 83 TNWICGMLHFTLLIK-----EPFLSLSSDILFEQGEARFTWRLAR-----GVISRD 131  
 XX 166 DSAILVEKESHPLEP-----CDT--LDNNEIVATIKF 197  
 XX 132 DEVFPEFQANSHFVKFKYADYDGTGRCVOLVKNLRILVKELVF 176

RESULT 13  
 ABG02357  
 ID ABG02357 standard; protein; 255 AA.  
 XX  
 XX AC ABG02357;  
 XX  
 XX DT 13-FEB-2002 (first entry)  
 XX  
 XX DE Novel human diagnostic protein #2348.  
 XX  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX PN WO200175067-A2.  
 XX  
 XX PD 11-OCT-2001.  
 XX  
 XX PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 XX PR 31-MAR-2000; 2000US-00540217.  
 XX 23-AUG-2000; 2000US-00649167.  
 XX  
 XX PA (HYSE-) HYSEQ INC.  
 XX  
 XX PI Dmanac RT, Liu C, Tang YT;  
 XX  
 XX DR WPI, 2001-639362/73.  
 XX DR N-PSDB; AAS66544.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 XX diagnostics, forensics, gene mapping, identification of mutations  
 XX responsible for genetic disorders or other traits and to assess  
 XX biodiversity.  
 XX  
 XX PS Claim 20; SEQ ID NO 32716; 103pp; English.  
 XX  
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 XX sequences. (I) is useful as hybridisation probes, polymerase chain  
 XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 XX and in recombinant production of (II). The polynucleotides are also used  
 XX in diagnostics as expressed sequence tags for identifying expressed  
 XX genes. (II) is useful in gene therapy techniques to restore normal  
 XX activity of (II) or to treat disease states involving (II). (II) is  
 XX useful for generating antibodies against it, detecting or quantitating a  
 XX polypeptide in tissue, as molecular weight markers and as a food  
 XX supplement. (II) and its binding partners are useful in medical imaging  
 XX of sites expressing (II). (I) and (II) are useful for treating disorders  
 XX involving aberrant protein expression or biological activity. The  
 XX polypeptide and polynucleotide sequences have applications in  
 XX diagnostics, forensics, gene mapping, identification of mutations  
 XX responsible for genetic disorders or other traits to assess biodiversity  
 XX and to produce other types of data and products dependent on DNA and  
 XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 XX amino acid sequences of the invention. Note: The sequence data for this  
 XX patent did not appear in the printed specification, but was obtained in  
 XX electronic format directly from WIPO at  
 XX ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 255 AA;  
 XX  
 XX Query Match 6.7%; Score 103; DB 4; Length 255;  
 XX Best Local Similarity 24.8%; Pred. No. 0.042;  
 XX Matches 41; Conservative 21; Mismatches 65; Indels 38; Gaps 8;  
 XX  
 XX 58 TAENAPAPETEDVSRNRYKVEFGMCVTGIGVR-EVILNCGCGSGKCVVVEGR 116  
 XX 25 TPKTALPEKIOEAVGKVI--INATTCVTGIGLKEETVCEVGPDRKCKQTRLECL 82  
 XX 117 GPTDCG-----GKPISESLSEVRLACIHTSP---NRFKYMVKLLRDOOSIILVN 165



DB 83 TNWICGMHFTLLIGK-----EFELSSCLSSDILFPGQAFRTFTWRLAR-----GVISTD 131  
 QY 166 DSALIEVKESHPLAFB-----CDT--LDNNEIVATIKF 197  
 DB 132 DEVFKEPQANSHFVFKYAOEYDSGTYRCDVQLVNKLRLVKLYLF 176  
 RESULT 14  
 AAU96170 ID AAU96170 standard; protein; 255 AA.  
 AAU96170;  
 02-JUL-2002 (first entry)  
 Human secreted protein, SEQ ID NO 72.  
 Human; secreted protein; autoimmune disease; rheumatoid arthritis;  
 hyperproliferative disorder; neoplasm; breast; liver; ischemia;  
 cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
 angiodysplasia; nervous system disorder; Alzheimer's disease; infection;  
 corneal infection; wound healing; ocular disorder; skin aging; sunburn;  
 epithelial cell proliferation; organ transplantation; food additive;  
 food storage.  
 Homo sapiens.  
 WO200224721-A1.  
 28-MAR-2002.  
 09-JAN-2001; 2001WO-US000544.  
 20-SEP-2000; 2000US-0234211P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Komatsu GA, Baker KP, Rosen CA, Birse CE, Soppet DR,  
 Olsen HS, Ni J, Fiascello M, Moore PA, Wei P, Ebner R, Duan DR;  
 Shi Y, Choi GH;  
 MPI; 2002-330012/36.  
 N-PSDB; ABK69091.  
 Isolated nucleic acid molecule encoding a human secreted protein is used  
 in preventing, treating or ameliorating a medical condition.  
 Claim 11; Page 498-499; 562pp; English.  
 The invention relates to an isolated nucleic acid molecule (I) encoding a  
 human secreted protein (II). (I) and (II) are used to prevent, treat or  
 ameliorate a medical condition in e.g. humans, mice, rabbits, goats,  
 horses, cats, dogs, chickens or sheep. (I) and (II) are also used in  
 diagnosing a pathological condition or susceptibility to a pathological  
 condition. The antibodies to (II) can also be used in alleviating  
 symptoms associated with the disorders and in diagnostic immunoassays  
 e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).  
 Disorders which are diagnosed or treated include autoimmune diseases e.g.  
 rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the  
 breast or liver, cardiovascular disorders e.g. cardiac arrest,  
 cerebrovascular disorders e.g. cerebral ischemia, angiodysplasia, nervous  
 system disorders e.g. Alzheimer's disease, infections caused by bacteria,  
 viruses and fungi and ocular disorders e.g. corneal infection. The  
 polypeptides can also be used to aid wound healing and epithelial cell  
 proliferation, to prevent skin aging due to sunburn, to maintain organs  
 before transplantation, for supporting cell culture of primary tissues,  
 to regenerate tissues and in chemotaxis. The polypeptides can also be  
 used as a food additive or preservative to increase or decrease storage  
 capabilities. AAU96165-AAU96237 represent human secreted protein  
 sequences and related sequences used in expression of the secreted  
 proteins as described in examples of the invention  
 Sequence 255 AA;

Query Match 6.7%; Score 103; DB 5; Length 255;  
 Best Local Similarity 24.8%; Pred. No. 0.042;  
 Matches 41; Conservative 21; Mismatches 65; Indels 36; Gaps 8;  
 QY 58 TAENYAPPEPDSVSNRYVKEVRFQMCTVTCGIGVR-EVILITNGCGESKCVAVVERCR 116  
 DB 25 TPKTLAIPKQOENAGKVI--INATTCVTCGLGYKEETVCVGPDPVRRKQOTRLBCL 82  
 QY 117 GPTDCG-----GKPISESLSEVRLACHTSP-----NPKWKKLLPDQGSIIILVN 165  
 DB 83 TNWICGMHFTLLIGK-----EFELSSCLSSDILFPGQAFRTFTWRLAR-----GVISTD 131  
 QY 166 DSALIEVKESHPLAFB-----CDT--LDNNEIVATIKF 197  
 DB 132 DEVFKEPQANSHFVFKYAOEYDSGTYRCDVQLVNKLRLVKLYLF 176  
 RESULT 15  
 ABG64819 ID ABG64819 standard; protein; 255 AA.  
 ABG64819;  
 27-AUG-2002 (first entry)  
 Human albumin fusion protein #1494.  
 Albumin fusion protein; therapeutic protein X; human albumin; HA;  
 human serum albumin; HSA; cancer; reproductive disorder;  
 digestive disorder; immune disorder; endocrine disorder;  
 haematopoietic disorder; neural disorder; connective disorder;  
 cytostatic; antifertility; antiinflammatory; anticancer;  
 immunomodulator; anti-HIV; antidiabetic; haemostatic; nocotropic;  
 neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
 osteopathic; antiarthritic.  
 Homo sapiens.  
 Synthetic.  
 WO200177137-A1.  
 18-OCT-2001.  
 12-APR-2001; 2001WO-US011988.  
 12-APR-2000; 2000US-0229358P.  
 25-APR-2000; 2000US-0199384P.  
 21-DEC-2000; 2000US-0256931P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Rosen CA, Haseltine WA;  
 MPI; 2002-010886/01.  
 New fusion protein for treating disease e.g. diabetes comprises an  
 albumin fused to a therapeutic protein.  
 Claim 1; Page 1522-1523; 2102pp; English.  
 The present invention relates to albumin fusion proteins comprising a  
 therapeutic protein X and human albumin (HA), also known as human serum  
 albumin, (HSA). The proteins are useful for treating a disease or disorder  
 that may be modulated by therapeutic protein X. The albumin extends the  
 shelf-life of protein X, and may increase its biological in vitro/in vivo  
 activity. The protein is useful for treating and diagnosing disorders  
 such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's  
 disease, ulcerative colitis), immune disorders (e.g. acquired  
 immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),  
 haematopoietic disorders, neural disorders (e.g. Alzheimer's,  
 Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,  
 schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).

CC ABG63326-ABG65518 represent albumin fusion proteins of the invention  
 XX Sequence 255 AA;

Query Match 6.7%; Score 103; DB 5; Length 255;  
 Best Local Similarity 24.8%; Pred. No. 0.042;  
 Matches 41; Conservative 21; Mismatches 65; Indels 38; Gaps 8;

QY	58	TAENYAPPEDEDVSNRNVKVEVEFGMCTVTGIGIVR-EVILINGCPGSGSKCVAVVEECR	116
DB	25	TPKTLAIPEKLEQBAVGKVI--INATTCVTTCGLGYKESTVCEVGPDPGVRRKQOTRECL	82
QY	117	GPTDQGW-----GKPISESLSEVRLACIHHSPL---NRKVMKLLRQDQSIILVN	165
DB	83	TNWICGMHLFTILICK-----EPRLSCLSDDLLEFGQBAFRFTWRLAR-----GVISTD	131
QY	166	DSAILLEVRKESHPLAFE-----CDT--LDNNEIVATIKF	197
DB	132	DEVFKPFOANSHFVFKYQAQSYDSCTYRCDVQLVKNLRIVKRLYF	176

Search completed: April 15, 2005, 14:11:58  
 Job time : 175 secs



N/Alternate names: kexin homolog; paired-basic endopeptidase; prohormone-processing endo  
 C/Species: Bos primigenius laurus (cattle)  
 C/Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
 C/Accession: I46044; S41191  
 R/Vey, M.; Schaefer, W.; Berghofer, S.; Klenk, H.D.; Garten, W.  
 J. Cell Biol. 127, 1829-1842, 1994  
 A/Title: Maturation of the trans-Golgi network protease furin: compartmentalization of f  
 A/Reference number: A55189; MUID:95105228; PMID:7806563  
 A/Accession: I46044  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-797 <VEY>  
 A/Cross-references: UNIPROT:Q28193; EMBL:X75956; NID:g439648; PIDN:CAA53569.1; PID:g4396  
 C/Function:  
 A/Description: cleavage of precursor proteins during constitutive secretory pathway at R  
 C/Superfamily: kexin; subtilisin homology  
 C/Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein  
 F:1-26/Domain: signal sequence #status predicted <SIG>  
 F:27-797/Product: furin #status predicted <MAT>  
 F:27-718/Domain: extracellular #status predicted <EXT>  
 F:144-382/Domain: subtilisin homology <SRT>  
 F:719-739/Domain: transmembrane #status predicted <TM>  
 F:740-797/Domain: intracellular #status predicted <INT>  
 F:153,194,368/Active site: Asp, His, Ser #status predicted  
 F:387,440,553/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.1%; Score 92.5; DB 1; Length 797;  
 Best Local Similarity 21.7%; Pred. No. 7.8;  
 Matches 76; Conservative 32; Mismatches 103; Indels 139; Gaps 18;

1 MSPRGTGSGAGLMTVGMILLAGL---QSARGTNTAIVQDAGLAHBEGBEETENNDS 57  
 VSPGKTRST-----LLAARPHDYSADGFNDWAFMTTHSDWEDSDSGEWTLEIENIS 555  
 58 TAENV-----APPE-----TEDVSNRVNVEVERGM-----CTVTCG 89  
 556 EANNVGTITKFTLVLYGTAPBGLPFPPEISIGCKTILTSQACVCEBGSFLHOKNCVQICP 615  
 90 IGRV-----EVLITNGCPGSEKCVVRVEBCRG--PTDCGKGPISBSLES 133  
 616 PGFAPQVLDTHYSTENDVEIIRASVCTPCASCAT---CQGPAPTDC-LSCPSHASIDP 670  
 134 VRLACIHSPNLPFRYMKLRQDQSGIILVNDASILVRKSH-----PLAFCD 184  
 671 VEQTC-----SRQSSS-----RSHHQQPPPPPPPPAP-- 700  
 185 TLDNNEIVATIKFTVYTSSELOMRSSLPATDALIFVLTIGVILCVPIIFLFIPIINW 244  
 701 -----VAT-----EPRLRADLPESHLPVVAAGISCAFIY---LVFTVFLVQL 741  
 245 AAVTAFWGAKASTPEVQSESSVRYSKSTSLDQLPTEPGEEDALSENNE 294  
 742 RSGFSFSGVAKVYT-----MDRGLISYKG-----LPPE-----AMQE 772

Db 742 RSGFSFSGVAKVYT-----MDRGLISYKG-----LPPE-----AMQE 772

QY 245 AAVTAFWGAKASTPEVQSESSVRYSKSTSLDQLPTEPGEEDALSENNE 294

Db 742 RSGFSFSGVAKVYT-----MDRGLISYKG-----LPPE-----AMQE 772

RESULT 3

A38198  
 A/Alternate names: alpha-1 chain, pancreatic - human  
 C/Species: Homo sapiens (man)  
 C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Jul-2000  
 C/Accession: A38198  
 R/Setno, S.; Chen, L.; Setno, M.; Blondel, O.; Takeda, J.; Johnson, J.H.; Bell, G.I.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 584-588, 1992  
 A/Title: Cloning of the alpha 1 subunit of a voltage-dependent calcium channel expressed  
 A/Reference number: A38198; MUID:92115705; PMID:1309948  
 A/Accession: A38198  
 A/Molecule type: mRNA  
 A/Residues: 1-2181 <SEI>  
 A/Cross-references: GB:M83566; NID:g179751; PIDN:AAA5629.1; PID:g179752  
 A/Experimental source: pancreatic beta cells  
 C/Superfamily: voltage-dependent calcium channel protein alpha-1 chain  
 C/Keywords: membrane protein; voltage-gated ion channel

Query Match 5.9%; Score 89.5; DB 2; Length 2181;  
 Best Local Similarity 21.2%; Pred. No. 43;  
 Matches 71; Conservative 46; Mismatches 97; Indels 121; Gaps 18;

32 VTAIVQDAGLAHBEGBEETENNDSFAENYAPPEDEVSNRVNVEVERGMCTVTCGIG 91  
 442 ITQA-EDIDNEBEGEBGRNTSMPTS-----ETESVNTENVSGBE----- 484  
 QY 92 VREVLITNGCPG-----GSEKCVVRVEBCRGPTDC-GWGPISBSLESVR----- 135  
 Db 485 -----NRGCCGSLMCWRRRGAAGA-----GPSGCRMGQALSKSKLSRRWRMNR 530  
 QY 136 -----LACIHSPNLPFRYMK---WTLNOD-QQSIILVNDASIL 170  
 Db 531 FNRRCRAVKSVTFWLVIVLPLNTLTSSSEHYNOPDWLTQIDANKVLLALFTCEM 590  
 QY 171 EVRKESHP-----AFEC-----DTLDNNEIVATIKFTVYTSSEL-----QMR 208  
 Db 591 LVQNVSLGLQVFPVSLRNPFCFVQGITFTILVELIEMSLGISVRCRLIRIFKYT 650  
 QY 209 R-SLEPATDALI-FVLTIGVILCVPIIFLFIPIINMAAVKAFWGAKASTPEVQSESS 265  
 Db 651 RMTSLSLVAVSLNMSKMSISALLLFLFIIFSLIG---MQLFGCKFNFDETQKRS 706  
 QY 266 SVRYKDSISDQLPTE-----MPGEDDALSEN 293  
 Db 707 -----TFDNFPOALLTFVQILLGED-----WN 728

RESULT 4

KKRT  
 furin (EC 3.4.21.75) precursor - rat  
 N/Alternate names: kexin homolog; paired-basic endopeptidase; prohormone-processing endo  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
 C/Accession: S13106  
 R/Misumi, Y.; Sohma, M.; Ikehara, Y.  
 Nucleic Acids Res. 18, 6719, 1990  
 A/Title: Sequence of the cDNA encoding rat furin, a possible propeptide-processing endop  
 A/Reference number: S13106; MUID:91067492; PMID:2251148  
 A/Accession: S13106  
 A/Molecule type: mRNA  
 A/Residues: 1-793 <MTS>  
 A/Cross-references: UNIPROT:P23377; EMBL:X55660; NID:g56171; PIDN:CAA3193.1; PID:g56172  
 C/Comment: This subtilisin-like endopeptidase removes paired basic residues to process f  
 C/Superfamily: kexin; subtilisin homology  
 C/Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein  
 F:1-26/Domain: signal sequence #status predicted <SIG>  
 F:27-793/Product: furin #status predicted <MAT>  
 F:27-714/Domain: extracellular #status predicted <EXT>  
 F:144-382/Domain: subtilisin homology <SRT>  
 F:715-735/Domain: transmembrane #status predicted <TM>  
 F:736-793/Domain: intracellular #status predicted <INT>  
 F:153,194,368/Active site: Asp, His, Ser #status predicted  
 F:387,440,553/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.7%; Score 87.5; DB 1; Length 793;  
 Best Local Similarity 21.7%; Pred. No. 21;  
 Matches 75; Conservative 30; Mismatches 106; Indels 135; Gaps 17;

1 MSPRGTGSGAGLMTVGMILLAGL---QSARGTNTAIVQDAGLAHBEGBEETENNDS 57  
 VSPGKTRST-----LLAARPHDYSADGFNDWAFMTTHSDWEDSDSGEWTLEIENIS 555  
 58 TAENV-----APPE-----TEDVSNRVNVEVERGM-----CTVTCG 89  
 556 EANNVGTITKFTLVLYGTAPBGLPFPPEISIGCKTILTSQACVCEBGSFLHOKNCVQICP 615  
 90 IGRV-----EVLITNGCPGSEKCVVRVEBCRG--PTDCGKGPISBSLES 133  
 616 PGFAPQVLDTHYSTENDVEIIRASVCTPCASCAT---CQGPAPTDC-LSCPSHASIDP 670

Query 134 VLACHTSPINFRKMKKLLRQDOQSIIIVNDSAILFVRKSHPLAECDTLNNNEIVA 193  
 Db 671 VEQTC-----SRQSOS-----RSRPOPPALRPV----- 698  
 Qy 194 TIKETVYTSSELQNR-----RSSLPATDAALFVLTIGVILCVFIIFLIPIINMAVK 248  
 Db 699 -----EVEPRLRAGLASHLPEVLAGLSCLI-----IALIFGVIFLFRCSGF 741  
 Qy 249 AFNGAKASTPEVQSSRSVKYDSTSLDQLPTEMGEDDALSNNNE 294  
 Db 742 SFRGVKVTY-----MDRGLISYKG-----LPEE-----AWQS 768

RESULT 5

SI2783  
 OX40 antigen precursor - rat  
 N/Alternate names: nerve growth factor receptor homolog  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C/Accession: S12783; S08036  
 R/Mallet, S.; Fossum, S.; Barclay, A.N.  
 EMBO J. 9, 1063-1068, 1990  
 A/Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lymphocyte  
 A/Reference number: S12783; MUID:90214614; PMID:2157591  
 A/Accession: S12783

A/Molecule type: mRNA  
 A/Residues: 1-271 <MAL>  
 A/Cross-references: UNIPROT:P15725; EMBL:X17037; NID:G57830; PID:CAA3497.1; PID:G57831  
 C/Superfamily: CD27 antigen; NGF receptor repeat homology  
 C/Keywords: growth factor receptor; transmembrane protein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-271/Product: signal sequence #status predicted <MAT>  
 F:211-235/Domain: transmembrane #status predicted <TM>

Query Match 5.7%; Score 86.5; DB 2; Length 271;  
 Best Local Similarity 21.1%; Pred. No. 7.7; Indels 53; Gaps 10;  
 Matches 42; Conservative 35; Mismatches 69;

Qy 99 NGCGESKCVVRVECGKPTDCGKGPISSELSVRLACT-----HSP-LNRKRMWK 152  
 Db 93 NCTPTEDTVC-----QCRPGT-----OPRODSHKLGVDCVCPGHPSPGNOACKPW- 141  
 Qy 153 LLRDQOQSIIIVNDSAILFVRKSHPLAECDTL-DNNEIVATIKR-----TVYTS 202  
 Db 142 -----TICTISGQIRHPASNSLDTVCEDNSLTLTMEYRTFRFTYPS 188  
 Qy 203 SELQWRSSLPATDA-----ALIFVLTIGVILCVFIIFLIPIINMAA-----VRAF 250  
 Db 189 TTWPRRTSQDLPSTPLVAPSGPAFAVILGLGLAPLTVLALVILRKAMRSPMPKPC 248  
 Qy 251 WGAKASTPEVQSSRSVKY 269  
 Db 249 WGNSEFRT-IOEQDTTHF 266

RESULT 6

A32068  
 circumsporozoite protein - Plasmodium vivax (strain North Korean) (fragment)  
 C/Species: Plasmodium vivax  
 C/Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 09-Jul-2004  
 C/Accession: A32068  
 R/Arnott, D.B.; Barnwell, J.W.; Stewart, M.J.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 8102-8106, 1988  
 A/Title: Does biased gene conversion influence polymorphism in the circumsporozoite prot  
 A/Reference number: A32068; MUID:89042133; PMID:3054880  
 A/Accession: A32068  
 A/Molecule type: DNA

A/Residues: 1-367 <AR>  
 A/Cross-references: UNIPROT:Q26169; GB:M0670; GB:J04090; NID:G160183; PID:AAA29534.1;  
 C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 F:293-345/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 5.6%; Score 85.5; DB 2; Length 367;

Best Local Similarity 30.5%; Pred. No. 13;  
 Matches 25; Conservative 8; Mismatches 28; Indels 21; Gaps 2;  
 Qy 22 AGLOSAGKTNTAIVQAGLAHGEGBEETNNSETAENYAPPTEDVSNRYKE--- 78  
 Db 244 AGGQAAGGNANANKAEADGAGNAGAGGQGNNEGANA-----PNEKSVKEYILD 292  
 Qy 79 -----VEFGKCTVTCGIGVR 93  
 Db 293 KYRATVGTETPTCSYTCGVGR 314

RESULT 7

OZ20AL  
 circumsporozoite protein precursor - Plasmodium cynomolgi (strain London)  
 N/Alternate names: major sporozoite surface antigen  
 C/Species: Plasmodium cynomolgi  
 C/Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 07-Nov-1997  
 C/Accession: A26255  
 R/Gallinek, M.R.; Arnott, D.E.; Cochran, A.H.; Barnwell, J.W.; Nussenzweig, R.S.; Enea, V  
 Cell 48, 311-319, 1987  
 A/Title: The circumsporozoite gene of the Plasmodium cynomolgi complex.  
 A/Reference number: A26255; MUID:87102878; PMID:3802196  
 A/Accession: A26255  
 A/Molecule type: DNA

A/Residues: 1-378 <GAL>  
 C/Comment: There are three distinct regions in the mature circumsporozoite protein, the e  
 obic membrane-anchoring sequence.  
 C/Comment: There are 19 tandem copies of a 6-residue repeat and 6 copies of an 11-residue  
 C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 C/Keywords: sporozoite; surface antigen; tandem repeat  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-378/Product: circumsporozoite protein #status predicted <MAT>  
 F:198-211/Region: 6-residue repeats  
 F:212-277/Region: 11-residue repeats  
 F:303-356/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 5.6%; Score 85; DB 1; Length 378;  
 Best Local Similarity 29.7%; Pred. No. 15;  
 Matches 38; Conservative 11; Mismatches 51; Indels 28; Gaps 4;

Qy 22 AGLOSAGKTNTAIVQAGLAHGEGBEETNNSETAENYAPPTEDVSNRYKE--- 78  
 Db 259 AGGQAAGGNAGAGGQAGAGAGGQ-----QNN-----EGANNPAKLVKEYILD 303  
 Qy 79 -----VEFGKCTVTCGIGVR---REVLTNGCGESKCVVRVECGKPTDCGKGPIS 128  
 Db 304 KIRSTIGVEMSPCSYTCGKGVRRKRVANAKKEBELDVNDLLETVCCTDKCAGIFNVVS 363  
 Qy 129 ESLSESVRL 136  
 Db 364 NSLGLVIL 371

RESULT 8

C41156  
 circumsporozoite protein - Plasmodium vivax (isolate B7-4)  
 C/Species: Plasmodium vivax  
 C/Date: 05-Jun-1992 #sequence\_revision 05-Jun-1992 #text\_change 09-Jul-2004  
 C/Accession: C41156  
 R/Gari, S.H.; Goldman, I.F.; Povoa, M.M.; Oliveira, S.; Alperis, M.P.; Lal, A.A.  
 J. Biol. Chem. 266, 16297-16300, 1991  
 A/Title: Wide distribution of the variant form of the human malaria parasite Plasmodium v  
 A/Reference number: A41156; MUID:91358402; PMID:1885563  
 A/Accession: C41156  
 A/Status: Preliminary  
 A/Molecule type: DNA

A/Residues: 1-387 <QAR>  
 A/Cross-references: UNIPROT:Q7M3W8; GB:M69062  
 C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
 F:312-365/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 5.5%; Score 84.5; DB 2; Length 387;

Best Local Similarity 31.4%; Pred. No. 17; Matches 38; Conservative 15; Mismatches 51; Indels 17; Gaps 5;

QY 22 AGLOSARGTNTAAVODAGLAHGESEETENNDSETAENYAPPETEDVSNRYNKE--- 78  
Db 271 AGGQAAGG---NAANKKAGDAGAGG---QNNEGAN---APNEKSVTEYLDKVRATVG 319

QY 79 VEFMCCTVTCGIGV---REVLTNCGCGESKCVVRVEECGPTDCGKGTISLSLSYR 135  
Db 320 TEMTPCVCVTCGVVRVRVRVATNKKPEDLTLDLETCTVCTMDKCAGIFNVVNSSLGLVI 379

QY 136 L 136  
Db 380 L 380

RESULT 9  
A41156  
circumsporozoite protein - Plasmodium vivax (isolate P19/D)  
C/Species: Plasmodium vivax  
C/Date: 05-Jun-1992 #sequence\_revision 05-Jun-1992 #text\_change 09-Jul-2004  
C/Accession: A41156  
R/Orig: S.H.; Goldman, I.F.; Povoa, M.M.; Oliveira, S.; Alperis, M.P.; Lal, A.A.  
J. Biol. Chem. 266, 16297-16300, 1991  
A/Title: Wide distribution of the variant form of the human malaria parasite Plasmodium  
A/Reference number: A41156; MUID:91358402; PMID:1885563  
A/Accession: A41156  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-395 <OAR>  
A/Cross-references: UNIPROT:Q7M3X0; GB:M69059  
C/Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
P/320-373/Domain: thrombospondin type 1 repeat homology <TR1>

Query Match 5.5%; Score 84; DB 2; Length 395;  
Best Local Similarity 34.1%; Pred. No. 19;  
Matches 28; Conservative 9; Mismatches 17; Indels 28; Gaps 4;

QY 22 AGLOSARGTNTAAVODAGLAHGESEETENNDSETAENYAPPETEDVSNRYNKE--- 78  
Db 279 AGGQAAGG---NAANKKAGDAGAGG---QNNEGANA-----TNEKSVTEYLD 320

QY 79 -----VEFGMCTVTCGIGVR 93  
Db 321 KVRATVGTETPCTCGVVR 342

RESULT 10  
I48771  
SLP(W7) - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C/Accession: I48771  
R/Hemenway, C.; Kallf, M.; Stavenhagen, J.; Walthall, D.; Robins, D.  
Nucleic Acids Res. 14, 2539-2554, 1986  
A/Title: Sequence comparison of alleles of the fourth component of complement (C4) and B  
A/Reference number: I48774; MUID:86176748; PMID:3008092  
A/Accession: I48771  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-594 <RBS>  
A/Cross-references: UNIPROT:Q62238; EMBL:X06454; NID:g54105; PIDN:CAA29760.1; PID:g54106  
C/Superfamily: alpha-2-macroglobulin

Query Match 5.5%; Score 84; DB 2; Length 594;  
Best Local Similarity 22.9%; Pred. No. 30;  
Matches 50; Conservative 31; Mismatches 77; Indels 60; Gaps 11;

QY 20 LLAGQSARG-TNTVTAAYODAGLAH-EGEGSEETENNDSETAENYAPPETEDVSNRYNKE 77  
Db 340 LLAGSHALRGDLKLTSLSDRYVSHFETDGPVLLYFDS-----VPTTRCVGFGASQ 392

QY 78 EVERGCM-----CTVTCGIGVREVLTNCGCGESKCVVRVEECGPTD 120

Db 393 EVVAVGLVOPSSAVLYDYPSDHCVSFPYAPTKSQLATLQSG-----DVCQ----- 439

QY 121 CGWCK--PISESL-----SVRLACTHTSPLEFNFKYMKLLRDQOSIILVNDSDAI 169  
Db 440 CAQCKCPPLRLSLERRVVDKGYMRPACYA--PRVYGFVVKYLRBDGRAAPRLPESKI 497

QY 170 LEVAKESHPLAFEDTLDDNNEIVATIKFTVYTSSELQM 207  
Db 498 TVV-----LHFRTDVMAS--IGQTRNPLRTSGRLK 527

RESULT 11  
T09986  
Probable transcription termination factor - Mycobacterium leprae  
C/Species: Mycobacterium leprae  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: T09986  
R/Robison, K.  
submitted to the EMBL Data Library, September 1994  
A/Reference number: Z16911  
A/Accession: T09986  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-610 <ROB>  
A/Cross-references: UNIPROT:P45835; EMBL:U15186; NID:g699323; PID:g699332  
C/Keywords: ATP; transcription termination

Query Match 5.5%; Score 83.5; DB 2; Length 610;  
Best Local Similarity 24.1%; Pred. No. 34;  
Matches 60; Conservative 30; Mismatches 102; Indels 57; Gaps 14;

QY 4 RGTCGSKGLMTVWTL---LLAGQSARG-TNTVTAAYODAGLAHGESEB-ETENNSET 58  
Db 50 PALANQGVKGTSTMSRKSEILALAEBCRGQANGSVNDGPERDHGSAATALSTALAAQE 109

QY 59 AENTAPPETEDVSNRYNRYVEFGMCTVTCGIGVREVLT-----TNGCPGSEKSC 108  
Db 110 KONTAIVYVSRERRRGASREAD-----VTAGTSTAEATSDCGSTADDTTLOGGQSD- 163

QY 109 VVRVEBCRGPTDCGKGPISSELSVRLACTHTSPL-----NFKYMKLLR 155  
Db 164 -TKTER-RGF-DVNDGQVGEQSSSLQ-----PRGDDDEGRGGRGRFRDRDRRR 213

QY 156 QDQOSIILVNDSDALFRKES--HPLAFEDTLDDNNEIVATIKR-----TYTSSSELQMR 208  
Db 214 GERS-----GDGAERLRQDDVQVAGIILVDVNTYAFVRTSGYLAGPHDYVSMW-VR 267

QY 209 RSSLPATDA 217  
Db 268 KNGLRGDA 276

RESULT 12  
D75591  
Probable cation transporter - Deinococcus radiodurans (strain R1)  
C/Species: Deinococcus radiodurans  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: D75591  
R/White, O.; Bisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.D.; Lam, P.; McDonald, L.; Ueberbach, T.; Zalewski, C.; Mal  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A/Reference number: A75250; MUID:20036896; PMID:10567266  
A/Accession: D75591  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-375 <WHI>  
A/Cross-references: UNIPROT:Q9RYF7; GB:AB001863; GB:AB001825; NID:g6460670; PIDN:AAF12426  
A/Experimental source: strain R1  
C/Genetics:  
A/Gene: DRA0361

A/Map position: 2  
C/Superfamily: magnesium and cobalt transport protein

Query Match	5.4%;	Score 83;	DB 2;	Length 375;
Best Local Similarity	24.7%;	Pred. No. 22;		

Dy  
116 RGFPGDGGWGPRISE-SLESTVLACHTSPINRPFKY-----MMKTLRQDDQSII---LVN 165  
          ::               ::               :  
Db 24 RPPTDAGPDPPLDDSGGRAPNAHAAVFRDGKVRRHLPLDLNLARGLLDDPQAQFVWFDPVLN 83

```

Cy      166 -DSALIERKES--HPLAFECTLDNNEIVA-----TIKFTYTSSELQMRSSLPATD 216
      |||  : |||||  : |||  : |||  :
Db      84 FDPADLEALREFDHLPLAIF-DALVGGGRVKRKYPTDTFVFV-----LHGASLPQD 135

```

Dy 217 AALI--FVLTIGVILCVPIIFLFI---IINMAVKAPWKGKGS----- 256  
          |   |      |   |   |   |   |   |   |   |   |   |  
Db 136 RLQLHEMALFIGQRFVLTLIQHPLFADEEIVERMQLVPAMSSASGLYYIIIDTVDMI 195  
          |   |   |   |   |   |   |   |   |   |   |   |

286 DDA 288

Db	256	BEA	258
RESULT 13			

glu-rRNA amidotransferase, subunit A (gluA) homolog - Lyme disease spirochete  
C/Species: Borrelia burgdorferi (Lyme disease spirochete)  
C/Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 09-Jul-2004  
C/Accession: E70142

son, D.; Peterson, J.; Kesteven, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, I.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hattab, S.; Smith, H.O.; Venter, J.C.; Ambrose, S.M. 1997

A1:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*..  
A1:Reference number: A70100; MUID:98065943; PMID:9403665  
A1:Accession: E70142  
A1:Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecular type: DNA  
A: Residues: 1-496 <KLE>  
A: Cross-references: UNIPROT:O51317; GB:AE001140; GB:AE000783; NID:G2668233; PI  
A: Experimental source: strain B31

Query Match	5.4%;	Score 83;	DB 2;	Length 496;
Best Local Similarity	19.4%;	Pred. No. 30;		
Matches 43;	Conservative 44;	Mismatches 81;	Indels 54;	Gaps 8

```

OY 47 GBEETENDSEALANTAPETEDVSNRN--VVAEV-----EFGCTYTCGI 90
      | : : : : : | : | : | : |
Db 246 GSDKNDSTVDLPDDPYLKTESLQGRNLAVIKELSDLMNDKRVANSFAEFKDLLSKGI 305

```

Dy 91 GRRVILITNGCGEGBKCVRVBECKRPDCC-----GKPISSSLVSLKACH 140  
 ::||: ::||: |||||  
 Db 306 NKEVSI-----EINFLISITYIISPVEASSNLARYGLCGKRISEGLS----- 351  
 ::||: ::||: |||||

192 VATTIKETTYTSRRIOWRRSSLPATDAATIEVLITIGVILCVPI 233

Db 407 PESCPIITPTSFVXKFRUGLDDDDVKMTYSBICVIANLI 448  
RESULT 14

N.Alternate names: protein sir1844; uvra protein  
M.Contains: excision endonuclease ABC (EC 3.1.--.-) chain A  
C.Species: *Synechocystis* sp.  
A.Variety: PCC 6803

C;Accession: S77349  
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamiu, E.; Nakamura, Y.; Miyajima, N.  
O, K.; Okumura, S.; Shingo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
C. N. Res. 3, 109-136, 1996

A: Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. PCC 6803  
S: 5.  
A: Reference number: S74322; MUID: 97061201; PMID: 8905231  
A: Accession: S77349  
S: 5.1  
A: 5.1 sequence not shown  
S: 5.2  
A: 5.2 sequence not shown

A: Molecule type: DNA  
A: Residues: 1-970 <KAN>  
A: Cross-references: UNIPROT: P73412; EMBL: D90906; GE: AB001339; NID: g1652492; PIDD: BAAL74  
A: Note: The nucleotide sequence was submitted to the EMBL Data library, June 1996

C,Genetics: A,Gene: uvrA C,Function: has ATPase and DNA binding activity; involved in DNA repair

C;Keywords: ATP; DNA binding; DNA repair; duplication; Myotubularin; nucleotide binding; E;34-41/Region: nucleotide-binding motif A (P-loop)  
F;652-935/Domain: ATP-binding cassette homology <ABCB>  
F;669-676/Region: nucleotide-binding motif A (P-loop)

Query Match 5.4%; Score 83; DB 2; Length 970;  
 Best Local Similarity 25.6%; Pred. No. 63;  
 Matches 42; Conservative 21; Mismatches 65; Indels 36; Gaps 6;

```

Cy  24  LQ$ARGINV-----TAAVQDAGLARGGEERTENNDSPTA-----ENVAPPE  66
      | : | | | | | | | | | | | | | | | | | | | | | | | |
Db  228  LKQ$BGTALIDILDKPFLAVLDGSKODKEALKKAENGQAVHAPLPRKIIFSENFACBEH  287

```

09 6/-TIEDVSNKVVNAEAFGATCTGAGTG-VAEVLLINCGFEGSALVRWEDKSGFLGG 122  
:::|::|::|::|:  
Db 288 GAVNDELSPRLFNSPYGACPDCHGIGFVRSFCPDLVIPPEKPVYAI----- 337

[illegible]

A54504  
circumsporozoite protein - Plasmodium malariae (strain Uganda 1-CDC)  
C.Species: Plasmodium malariae  
C.Date: 06-Oct-1994 #sequence\_revision 04-Nov-1994 #text\_change 09-Jul-2004

C/Accession: A54304  
Rital, A.A.; de la Cruz, V.F.; Campbell, G.H.; Procell, P.M.; Collins, W.E.; McCutchan, M.L. Biochem. Parasitol. 30, 291-294, 1988  
A>Title: Structure of the circumsporozoite gene of *Plasmodium malariae*.  
A/Reference number: 254504. MIM: 60040073. DMT: 3054537

A1:Accession: A54504  
A1:Molecule type: DNA  
A1:Residues: 1-429 <LML>  
A1:Cross-references: UNIPROT:P13815; GB:J03992; NID:g160220; PIDN:AAA29557.1; PID:g16022

C:Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology  
C:Keywords: tandem repeat  
P:354-407/Domains: thrombospondin type 1 repeat homology <THRI>

42 AHRRGGRRRTNNNSRTARRVAPDPTEDVSN-----RVVVKEVVEGMCCTVCGIGVR 93  
Query Match 5.74% Score 24.07 DB 21 Domain 227  
Best Local Similarity 35.1% Pred.No. 28;  
Matches 20; Conservative 9; Mismatches 17; Indels 11; Gaps 3;

Job time : 46 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2005, 14:04:30 ; Search time 187 seconds  
(without alignments)  
805.087 Million cell updates/sec

Title: US-10-809-655-9  
Perfect score: 1528  
Sequence: 1 NSPRGCGCAGLMTVGNLL.....LDQLPTMGEDDALSNNNE 294

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1538	100.0	294	2	Q9HBV2	Q9HBV2 homo sapien
2	1115.5	73.0	305	2	Q9DA48	Q9DA48 mus musculu
3	311	20.4	90	2	Q80VM2	Q80VM2 mus musculu
4	110	7.2	1331	2	Q9HBD0	Q9HBD0 neurospora
5	104	6.8	1081	2	Q9UG31	Q9UG31 dirosophila
6	103	6.7	255	2	Q6P7N7	Q6P7N7 homo sapien
7	102	6.7	1091	2	Q7YU67	Q7YU67 dirosophila
8	102	6.7	1093	2	Q9VTT0	Q9VTT0 dirosophila
9	100	6.5	314	2	Q8IHMS	Q8IHMS plasmodium
10	99	6.5	255	2	Q6UVZ4	Q6UVZ4 homo sapien
11	99	6.5	628	2	Q73059	Q73059 treponema d
12	97.5	6.4	456	2	Q9FCV5	Q9FCV5 escherichia
13	97	6.3	709	1	KRPI_SCHPO	KRPI_SCHPO schistosac
14	95	6.2	1014	2	Q9VQ99	Q9VQ99 dirosophila
15	95	6.2	1353	2	Q9V907	Q9V907 dirosophila
16	94	6.1	259	2	Q6AXK1	Q6AXK1 m mus muscu
17	93.5	6.1	262	2	Q6AXW8	Q6AXW8 rattus norv
18	93	6.1	794	2	Q60426	Q60426 cricetus
19	92.5	6.1	797	1	FURI_BOVIN	FURI_BOVIN bos taurus
20	92	6.0	1949	2	Q9DF53	Q9DF53 brachydanio
21	91	6.0	401	1	CSP_PLACG	CSP_PLACG plasmodium
22	91	6.0	1213	2	Q6NZL6	Q6NZL6 mus musculu
23	90.5	5.9	1003	2	Q8TNN1	Q8TNN1 methanosaar
24	90.5	5.9	1643	2	Q91WX9	Q91WX9 rattus norv
25	90.5	5.9	2179	2	Q7TSP2	Q7TSP2 mus musculu
26	90	5.9	985	2	Q70G88	Q70G88 anopheles g
27	89.5	5.9	745	2	Q7P1Y6	Q7P1Y6 anopheles g
28	89.5	5.9	800	2	Q00835	Q00835 dictyostell
29	89	5.8	226	2	Q26136	Q26136 plasmodium
30	89	5.8	373	2	Q6PU93	Q6PU93 candida gla
31	89	5.8	378	1	CSP_PLACT	CSP_PLACT plasmodium

32	88.5	5.8	533	2	Q7XAF7	Q7XAF7 brassica na
33	88.5	5.8	1327	2	Q6PPA4	Q6PPA4 leishmania
34	88	5.8	368	2	Q8MPK1	Q8MPK1 plasmodium
35	88	5.8	456	2	Q659Z2	Q659Z2 escherichia
36	88	5.8	1067	2	Q84QV5	Q84QV5 oryza sativ
37	87.5	5.7	793	1	FURI_RAT	FURI_RAT rattus norv
38	87.5	5.7	1098	2	Q6BH26	Q6BH26 debaromyce
39	87	5.7	1020	2	Q6G918	Q6G918 uncultured
40	87	5.7	1047	2	Q6J155	Q6J155 ipomoea tri
41	86.5	5.7	271	1	TNR4_RAT	TNR4_RAT rattus norv
42	86.5	5.7	302	2	Q9GQL9	Q9GQL9 plasmodium
43	86.5	5.7	350	2	Q26117	Q26117 plasmodium
44	86.5	5.7	350	2	Q26118	Q26118 plasmodium
45	86.5	5.7	377	2	Q26116	Q26116 plasmodium

## ALIGNMENTS

## RESULT 1

ID Q9HBV2 PRELIMINARY; PRT; 294 AA.  
AC Q9HBV2;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Sperm acrosome membrane protein SAMP32 (Sperm acrosome associated  
DE 1).  
GN Name=SAMP32; Synonyms=SPACAL;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Butiria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21858310; PubMed=11870081; PubMed=11870081; Sen B.,  
RA Hao Z., Wolkowicz M.J., Shetty J., Klotz K., Bolling L., Sen B.,  
RA Westbrook V.A., Connor S., Flickinger C.J., Herr J.C.;  
RT "SAMP32, a testis-specific, isocitrogenic sperm acrosomal membrane-  
RT associated protein."  
RT [1]  
RT [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton B., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Rodriguez A.C., Grimwood J.W., Green B.D., Dickson M.C.,  
RA Krzyzanski M.I., Skaleka U., Smailus D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL [1]  
RL [1]  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=16673574442002;  
RA Strauberg R.;  
RA Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF203447; AAG31422.1;  
DR EMBL; BC029488; AAH29488.1;  
DR Genew; HGNC:14967; SPACAL.

DR GO:0008083; F: growth factor activity; IEA.  
 DR InterPro: IPR000762; PTN MK.  
 DR SEQUENCE 294 AA; 32143 MW; 60DB3107EB03D12 CRC64;  
 Query Match 100.0%; Score 1528; DB 2; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-122;  
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSRRGTGCGAGLMTVGMILLAGLQSGARGTNTAIVQDAGLAHGGEGEETENDESYA 60  
 DB 1 MSRRGTGCGAGLMTVGMILLAGLQSGARGTNTAIVQDAGLAHGGEGEETENDESYA 60  
 QY 61 NTAPEPTEVSNRNVAKEVEFGMCCTVTCGIGVREVLITNCGPGSEKCVVRVECGRPD 120  
 DB 61 NTAPEPTEVSNRNVAKEVEFGMCCTVTCGIGVREVLITNCGPGSEKCVVRVECGRPD 120  
 QY 121 CGMGKISSELSLESLACIHTSPINFRFKYMKLLRQDQSIILVNDSAILEVAKESHPLA 180  
 DB 121 CGMGKISSELSLESLACIHTSPINFRFKYMKLLRQDQSIILVNDSAILEVAKESHPLA 180  
 QY 181 FEGCDTLDNNEIVATIKFTYTTSELQMRSSLPATDAALIFVLTIGVITCVPIIFLIFI 240  
 DB 181 FEGCDTLDNNEIVATIKFTYTTSELQMRSSLPATDAALIFVLTIGVITCVPIIFLIFI 240  
 QY 241 IINMAVKAFGAKASTPEVQSSSVRYKSTSLDQPTMPGDDALSENN 294  
 DB 241 IINMAVKAFGAKASTPEVQSSSVRYKSTSLDQPTMPGDDALSENN 294

RESULT 2  
 ID Q9DA48 PRELIMINARY; PRT; 305 AA.  
 AC Q9DA48;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DB Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700021A02 product:weakly similar to SPERM ACROSOME MEMBRANE PROTEIN SAMP32.  
 GN Name=430540J03Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 NC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carinci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN PANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 403:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RA The PANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carinci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT Prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carinci P., Kono H., Akiyama J., Nishi K., Kitajima T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arikawa T., Bono H., Carinci P., Fukuda S., Furushiki Y., Furuno M., Hasegaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kaubakawa T., Kato H., Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Munazaki R., Ono M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi P., Tanaka T., Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK06187; BAB24447.1;  
 DR MGD: MGI:1914902; 4930540J03Rik.  
 DR GO:0008083; F: growth factor activity; IEA.  
 DR InterPro: IPR000762; PTN MK.  
 DR SEQUENCE 305 AA; 33342 MW; 9D68F7322E2DB398 CRC64;  
 Query Match 73.0%; Score 1115.5; DB 2; Length 305;  
 Best Local Similarity 68.6%; Pred. No. 2.7e-87;  
 Matches 210; Conservative 42; Mismatches 41; Indels 13; Gaps 5;

QY 1 MSRRGTGCGAGLMTVGMILLAGLQSGARGTNTAIVQDAGLAHGGEGEETENDESYA 53  
 DB 1 MSRRGTGCGAGLMTVGMILLAGLQSGARGTNTAIVQDAGLAHGGEGEETENDESYA 59  
 QY 54 -NDSETAENYAPPER---EDVSNRNVAKEVEFGMCCTVTCGIGVREVLITNCGPGSEKCV 109  
 DB 60 ENRGEVPESETTAEDAEVEVQNRITVKEVEFGMCCTVTCGIGVREVLITNCGPGSEKCV 119  
 QY 110 VAVEBCRGPTDCMGKPISELSLESLACIHTSPINFRFKYMKLLRQDQSIILVNDSAI 169  
 DB 120 VAVEBCRGPTDCMGKPISELSLESLACIHTSPINFRFKYMKLLRQDQSIILVNDSAI 179  
 QY 170 LEVRKSHPLAECDTLDNNEIVATIKFTYTTSELQMRSSLPATDAALIFVLTIGVIT 229  
 DB 180 LEVTRIRPLAECDTLDNNEIVATIKFTYTTSELQMRSSLPATDAALIFVLTIGVIT 239  
 QY 230 CVPIIFLIFIINNAVKAFGAKASTPEVQSSSVRYKSTSLDQPTMPGDDALSENN 288  
 DB 240 CIEVIFVLIIFIINNAVKAFGAKASTPEVQSSSVRYKSTSLDQPTMPGDDALSENN 299  
 QY 289 LSENN 294  
 DB 300 LSENN 305

RESULT 3  
 ID Q80VM2 PRELIMINARY; PRT; 90 AA.  
 AC Q80VM2;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DR Similar to RIKEN cDNA 4930540L03 gene.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=olfactory epithelium;  
 RX MBLIN=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Buetow K.H., Scheer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin J.L., Hong L.,  
 RA Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Rahe J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalske U., Smallos D.R., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=olfactory epithelium;  
 RA Strausberg R., et al. (2003) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 SQ SEQUENCE 90 AA; 10159 MW; A049C5010A9E5CC6 CRC64;  
 QY Query Match 20.4%; Score 311; DB 2; Length 90;  
 DB Best Local Similarity 67.4%; Pred. No. 6.4e-19;  
 Matches 58; Conservative 16; Mismatches 12; Indels 0; Gaps 0;  
 QY 191 IVATIKPTVTSSSELOMRSSSLPATDAALIFVLTIGVITVIFLFIINMAVAF 250  
 DB 1 MWASVYFVTTNLEQMRSSSPDTDAVLVFLTIGVITVIFLFIINMAVAF 60  
 DB 61 WGSKTATETIOSELSMRCEODKIE 86  
 QY 251 WGAQASTPEVSGSSVRYKSTSD 276  
 DB 61 WGSKTATETIOSELSMRCEODKIE 86  
 RESULT 4  
 Q9HED0 PRELIMINARY; PRT; 1331 AA.  
 Q9HED0  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Related to neomycin resistance protein NE01.  
 GN Name=99H12.220;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 RN NCBI\_TaxID=5141;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte V., Aign V., Hohelsel J., Brandt P., Fatmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL651018; CAC18258.2; -

DR GO:0016021; C:integral to membrane; IEA.  
 DR GO:0005524; F:ATP binding; IEA.  
 DR GO:0016887; F:ATPase activity; IEA.  
 DR GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO:0016787; F:hydrolase activity; IEA.  
 DR GO:0006812; P:cation transport; IEA.  
 DR GO:0008152; P:metabolism; IEA.  
 DR InterPro: IPR001757; ATPase E1-E2.  
 DR InterPro: IPR005834; Dehalo-like hydro.  
 DR InterPro: IPR006539; P1pase.  
 DR Pfam: PF00702; Hydrolase, 1.  
 DR PRINTS: PRO0119; CATAPASB.  
 DR TIGRFAMs: TIGR01652; ATPase-P1pase.  
 DR TIGRFAMs: TIGR01494; ATPase P-type; 3.  
 DR PROSITE: PS00154; ATPase E1-E2; UNKNOWN 1.  
 SQ SEQUENCE 1331 AA; 147482 MW; 9FC644FC83820608 CRC64;  
 QY Query Match 7.2%; Score 110; DB 2; Length 1331;  
 DB Best Local Similarity 25.2%; Pred. No. 2.8;  
 Matches 61; Conservative 38; Mismatches 93; Indels 50; Gaps 14;  
 QY 41 LAHGEGERETNNDSERARVYA---PPEEDV---SRRNVKVEPFMCVTGIGIRE 94  
 DB 378 LAHSAANKETEEQVPAKEMLLLDHVDVDDVGGSSKNKTKESN----- 425  
 QY 95 VILNCGCPGGS-KCVAVRECRGPTDCG-----GKPISESL---ESVRLACIHSP--- 143  
 DB 426 ---NNGAGSSSGRTFRTDQDETID--WGLRALPSQSLSTBELVRLAVTAKPKR 480  
 QY 144 LNRFRYMKLL--RQDQSLVNDASILEYRKSHPLAFECDTLDNNEIVA---TIK 197  
 DB 481 VNEFVGTLELLPSRDVMSGAANPREGDVKK--AAPSLID-NTRAMANTVIAASNTLAV 537  
 QY 198 TVTSSSELOMRSSSLPATDAALIF---VLTIGVITVIFLFIINMAVAFMAK 254  
 DB 538 IVYGPQTRALSSTPSKSTGLLEYINSLTLLC---PLTFLSTLVALLBGFSTAK 593  
 QY 255 AS 256  
 DB 594 GN 595  
 RESULT 5  
 Q9U631 PRELIMINARY; PRT; 1081 AA.  
 Q9U631  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Semaphorin 5C.  
 GN Name=Sema-5C;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pezomyzota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Oregon-R;  
 RX MEDLINE=20171070; PubMed=10704872; DOI=10.1016/S0925-4773(99)00297-X;  
 RA Khare N., Faacetelli N., Darocha S., Chiquet-Ehrismann R.,  
 RA Baumgartner S.;  
 RT "Expression patterns of two new members of the Semaphorin family in  
 RT Drosophila suggest early functions during embryogenesis";  
 RL Mech. Dev. 91:393-397 (2000).  
 DR EMBL; AF198084; AAF04860.1; -  
 DR HSSP; P07996; 1LSL.  
 DR FlyBase; FBgn0028679; Sema-5C.  
 DR GO:0016020; C:membrane; IEA.  
 DR GO:0003824; F:catalytic activity; IEA.  
 DR GO:0008199; F:ferric iron binding; IEA.  
 DR GO:0004872; P:receptor activity; IEA.  
 DR GO:0006725; P:aromatic compound metabolism; IEA.

DR GO:0007275; P:development; IRA.  
 DR InterPro; IPR000183; Decarboxylase.  
 DR InterPro; IPR000627; Dioxigenase.  
 DR InterPro; IPR003659; Plexin-like.  
 DR InterPro; IPR002165; Plexin\_repeat.  
 DR InterPro; IPR001627; Sema.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01437; PSI\_1.  
 DR Pfam; PF01403; Sema; 1.  
 DR Pfam; PF00090; TSP\_1; 6.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SMO0423; PSI\_1.  
 DR SMART; SMO0630; Sema; 1.  
 DR SMART; SMO0209; TSP1; 6.  
 DR PROSITE; PS00879; ODR\_DC\_2\_2; UNKNOWN\_1.  
 DR PROSITE; PS50092; TSP1; 6.  
 DR SEQUENCE 1081 AA; 120435 MW; 561071C831C431D3 CRC64;  
 SO

Query Match 6.8%; Score 104; DB 2; Length 1081;  
 Best Local Similarity 22.1%; Pred. No. 7.2; Mismatches 103; Indels 80; Gaps 12;  
 Matches 62; Conservative 35;

DR 43 HEGGEETENNDSSTAENYAPPETEDVSNRVKVEFG-----MCTVTCIGVR 93  
 DB 806 HRGGSQ-----SRVCMHACPAEBOUSSNDNEVHEHGMCSMSACSVCGLDLR 859  
 QY 94 EVIITNGCPGG-----ESRC-VRVVECRGPTDCGKPISSLSBVLACIHT 141  
 DB 860 R-RTRRCLAGHDLRCGRALBEGCKEMVPCENFLGWS--ANSEWSSCSGDIRL----- 910  
 QY 142 SPLRPFKTMGLLRDQDSIIIVNDSALLEVAKESHPLAFEDDTLDNNEIVATIKFTY 201  
 DB 911 -----RRRCVLVEQ-----PSMBCRG-----AEFEKIACV 936

QY 202 SSEIOWRRSSLPATDALIFVLTIGVILCVFIIFLI---FIITMAVKAFWGAKASTP 258  
 DB 937 PNECEIOTATATPIYIFVGLFTVACCLATYFTKKRFLSLBRLNKTITTTTAFD 996

QY 259 EVQGEQSSVRYKDTSLDQLP-----EMPGEDALISEMN 293  
 DB 997 TYPNQYSSLPTKD--YDQRPKQSSFRMFAKTSMLGN 1034

RESULT 6  
 Q6P7N7 PRELIMINARY; PRT; 255 AA.  
 AC 06P7N7;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Similar to RIKEN CDNA 4930429020.  
 GN Name=LOC388730;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 OK [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Peripheral Nervous System;  
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Scheffen C.F., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schenck C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haleh F.,  
 RA Datchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.B.,  
 RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carinici P., Prange C.T.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,  
 RA Bosak S.A., McEwen P.J., McKernan K.D., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Peripheral Nervous System;  
 RC Strausberg R.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC061592; AAB61592.1;  
 DR InterPro; IPR007110; Ig-like.  
 DR PROSITE; PS50835; Ig\_LIKE; 1.  
 DR SEQUENCE 255 AA; 28468 MW; 27FDBBAPFA/EO21 CRC64;  
 SO

Query Match 6.7%; Score 103; DB 2; Length 255;  
 Best Local Similarity 24.8%; Pred. No. 1.5; Mismatches 65; Indels 38; Gaps 8;  
 Matches 41; Conservative 21;

QY 58 TAENYAPPETEDVSNRVKVEFGMCTVTCIGVR-EVITNGCPGSGSKCVRVVECR 116  
 DB 25 TPXTAIPEKIQEAVGKI--INATCTVTCIGLGYKSEVCEVGPDPVRKRCOTRLECL 82  
 QY 117 GPTQCGM-----GKPISSLSBVLACIHTSL-----NFKTMGLLRDQDSIIIVN 165  
 DB 83 TMTWICGMLHFTILGK-----EPFLSCSLSDILEPQGEAFRTWRLAR-----GVISTD 131

QY 166 DSALIEVRKESHPLAF-----CDT--LDNNEIVATIKF 197  
 DB 132 DEVKPFOANSHFVKFKYQGEYDSGTRCDVQVKNRLVRLVLF 176

RESULT 7  
 Q7YU67 PRELIMINARY; PRT; 1091 AA.  
 AC 07YU67;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE RE68041D.  
 GN Name=Sema-5c;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 OK [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Berkeley;  
 RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champagne M., Chavez C., Dorsett V., Dresnek D., Patran D., Frits E.,  
 RA George R., Gonzalez M., Guarin H., Kromolli B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Parasas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceoliker S.;  
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BT009966; AAQ22435.1;  
 DR HSSP; Q92854; 10LZ.  
 DR PLYBase; PB00028679; Sema-5c.  
 DR GO; GO:0016020; C:membrane; IRA.  
 DR GO; GO:0003824; F:catalytic activity; IRA.  
 DR GO; GO:0008199; F:ferrie iron binding; IRA.  
 DR GO; GO:0004872; P:receptor activity; IRA.  
 DR GO; GO:0006725; P:aromatic compound metabolism; IRA.  
 DR GO; GO:0007275; P:development; IRA.  
 DR InterPro; IPR000183; Decarboxylase.  
 DR InterPro; IPR000627; Dioxigenase.  
 DR InterPro; IPR003659; Plexin-like.  
 DR InterPro; IPR002165; Plexin\_repeat.  
 DR InterPro; IPR001627; Sema.

RA Kimmel B.E., Kodira C.D., Kratich C., Kravitz S., Kuip D., Lai Z.,  
RA Laško P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merilov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclet J.M.,  
RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of *Drosophila melanogaster*.",  
RL Science 287:2185-2195 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise B., Hodgson A.,  
RA George R.A., Hoskins R.A., Lawerty T., Muzny D.M., Nelson C.R.,  
RA Paclet J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
RT *melanogaster* euchromatic genome sequence.",  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise B., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.,  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin  
RL genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.B.,  
RA Smith C.D., Tudy J.L., Whitfield S.E., Bayraktaroglu I., Berman B.P.,  
RA Batencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.,  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review.",  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AEO03542; AAP49966.3; --  
DR HSP, P07956; ILSL.  
RN InAct; Q9VT01; --  
DR FlyBase; Fgmn0028679; Sema-5c.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0008199; F:ferrie iron binding; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0006725; P:alometric compound metabolism; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR InterPro; IPR000183; Dicarboxylae2.  
DR InterPro; IPR000627; Dicarboxylae.  
DR InterPro; IPR003659; Plectin-like

```

DR InterPro; IPR002165; Pfexln_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF00090; TSP; 1; 6.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00209; TSP; 6.
DR PROSITE; PS00879; ODR_DC_2_2; UNKNOWN_1.
DR PROSITE; PS50092; TSP1; 5.
SQ SEQUENCE 1093 AA; 121927 MW; 8A36FELA22PBIDBF CRG64;

Query Match 6.7%; Score 102; DB 2; Length 1093;
Best Local Similarity 21.8%; Pred. No. 11;
Matches 61; Conservative 37; Mismatches 102; Indels 80; Gaps 12;

QY 43 HEGGSEETENNDELTENAVAPPEDEDVSNRNVAKEVERF-----MCTVTCGIGVR 93
DB 818 HREGSQ-----SRVCNTHACPAEBOLSNLDELHSGMGSSEWSACSVTGGLRL 871
QY 94 EVILYNGCPG-----ESKC-VVRYBECGPTDCGKPISELESVRLACIHT 141
DB 872 R--RTRCLAGHDLCCGRALBEOKCEVPCEDPLGWS--AMSESSCSDGIRL----- 922
QY 142 SPLRPFYKMKLRLQDQOSIILVNDALILEVRKESHPLAFECPTLDNNEIVATIKFTVYT 201
DB 923 -----RHRCCLVRO-----PSMSECRG-----AEFEKIVCV 948
QY 202 SSEIOWRSSLPADNALIFLITGVIIICVPIIFLI-----FIINMAVKAFWGKXSTP 258
DB 949 PNEBETQTATATPIYIFGLFTVACCLATYRFTKRFPLSBEALNRTTITTSFDP 1008
QY 259 EVQSGSSVRYKSTSLDQLPT-----EMPGEDALISEMN 293
DB 1009 TYRQVYSLPTKD--YYDQRKROSSFFMPAKTSLNGNGN 1046

RESULT 9
Q8IHM5 PRELIMINARY; PRT; 314 AA.
AC Q8IHM5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE R1fin.
GN ORFNames=PF11_0515;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=36329;
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Pung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pahn A., Nelson K.B., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shalimov S.J., Sub B., Peterson J., Angiolini S.,
RA Petrea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairclough A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPhaden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carnucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL EMBL; AE014843; AAN36084.1;
DR InterPro; IPR011038; Calycin.
DR InterPro; IPR00673; R1fin.
DR InterPro; IPR002858; R1fin_STEVOR.
DR Pfam; PF02009; R1fin_STEVOR; 1.
DR TIGRfam; TIGR01477; R1FIN; 1.

```

```

SQ SEQUENCE 314 AA; 35249 MW; F2F02922448AD040 CRG64;

Query Match 6.5%; Score 100; DB 2; Length 314;
Best Local Similarity 22.4%; Pred. No. 3.6;
Matches 56; Conservative 36; Mismatches 92; Indels 66; Gaps 9;

QY 52 ENNDETAENAVAPPEDEDVSNRNVAKE-----VE-----FG 82
DB 64 ENFROTQSQPEEBEEMIRKRCQEQCDKIKIIVKIKIESPAKLEKGLRGFG 123
QY 83 MCTVTCIGVREVLITNGCGGSEKCVRYBECGPTDCGKPISELESV 134
DB 124 LGVVAASVGIIGPIIVN-----ELKTAALVAAQGTGEAGIDKALEVVISKYGVNKLXGV 178
QY 135 RLACIHTSPLARFYMMLKLRQDQOSIILVNDALILEVRKESHPLAFECPTLDNNEIV-- 192
DB 179 ALERKITS--NNFQNVAFYIOAIRNTVWCSA--EPIDIGLCPLKOSLNGVLEPTK 234
QY 193 -----ATIKVTYTSSELQWRSSLPATDALLIFLITGV--IICVPIIFLI 238
DB 235 SISASQKVVAADATEKATITVTKAEV-----SAERKTSVNLVYALVYIALIVLVWVIT 290
QY 239 FIINMAVX 248
DB 291 FLIRYRRKK 300

RESULT 10
Q6UVZ4 PRELIMINARY; PRT; 255 AA.
ID Q6UVZ4
AC Q6UVZ4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE KVA12788.
GN ORFNames=UMQ2788;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya B., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chu C., Crowley C., Currell B., Deneil B., Dowd P.,
RA Bacon D., Foster J., Grimaldi C., Gu Q., Hase P.E., Helens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robble B., Sanchez C., Schoenfeld J.,
RA Seasholtz S., Simons L., Singh J., Smith J., Stinson J., Vagte A.,
RA Vanden R., Watanabe C., Wleand D., Woods K., Xie M.H., Yamaura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment."
RL Genome Res. 13:2265-2270(2003).
DR EMBL; AY359081; AAO89440.1;
DR InterPro; IPR007110; IG-Like.
DR PROSITE; PS50835; IG_LIKE; 1.
SQ SEQUENCE 255 AA; 28440 MW; E7FEBAF968A1B95 CRG64;

Query Match 6.5%; Score 99; DB 2; Length 255;
Best Local Similarity 24.2%; Pred. No. 3.4;
Matches 40; Conservative 22; Mismatches 65; Indels 38; Gaps 8;

QY 58 TAENVAPPEDEDVSNRNVAKEFMCTVTCGIGVR-EVILYNGCPGSEKCVRYBECR 116
DB 25 TPKTALPEKQEAIVGKI--INATCTCTVCGIGKEVCEVGPDPVRRKCTQRLACL 82
QY 117 GPIDCGW-----GKPISSLESVRLACIHTSPL-----NRKFMVQLRQDQOSIILVN 165
DB 83 TNNICGMHFTLILGK-----EFELSCSSDLIEFGQEAERFTWMLAR-----GVISTD 131

```

QY 166 DSAILLEVRKSHPLAF-----CDT--LDNNEIVATIKF 197  
 DB 132 DEVRKPRQANSHFVKFKYAGYDSCTRYCDVQVGNLALVRLYF 176

## RESULT 11

Q73059 PRELIMINARY; PRT; 628 AA.  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 OS OrderedLocustNames=TDE0585;  
 GN Treponema denticola.  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 NC NCBI\_TaxID=158;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35405 / DSM 14222;  
 RX PubMed=15064399; DOI=10.1073/pnas.0307639101;  
 RA Sebahdri R., Myers G.S.A., Tetteilin H., Eissen J.A., Heidelberg J.F.,  
 RA Dodson R.J., Davidson T.M., Deboy R.T., Fouts D.E., Haft D.H.,  
 RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,  
 RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,  
 RA Gheorgheorgis E., Geer K., Tsagaye G., Malek J.A., Ayodeji B.,  
 RA Shacterman S., McLeod M.P., Smajs D., Howell J.K., Pal S., Amin A.,  
 RA Vashisth P., McNeil T.Z., Xiang Q., Sodergren E., Baca E.,  
 RA Weinstock G.M., Norris S.J., Fraser C.M., Paulsen I.T.;  
 RA "Comparison of the genome of the oral pathogen Treponema denticola  
 RT with other spirochete genomes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651 (2004).  
 DR EMBL; AB017248; AAS11080.1; --.  
 DR TIGR; TDR0585; --.  
 DR Complete proteome.  
 SQ SEQUENCE 628 AA; 72538 MW; 8091B1A1BB9AAB42 CRC64;

Query Match 6.5%; Score 99; DB 2; Length 628;  
 Best Local Similarity 21.1%; Pred. No. 10; Indels 98; Gaps 10;

Matches 64; Conservative 35; Mismatches 106; Indels 98; Gaps 10;  
 QY 32 VTAAVQDAGLAHEGSEETENN---DSEIENYAPPEDEVSNNVYKEVEGACTVT 87  
 DB 307 ITQTDSDGVLGLGYSEITLONFMKXASSEKIIID-----ILFT 350  
 QY 88 CGIGREYILTNGCGESKCVRYVEBCRGFTDCMGKPISESLSVRLACHTSPILRF 147  
 DB 351 NSSNDRFYLIT-----EKVPLILSL-----INERKRVRLCT----- 384  
 QY 148 KYMKKLLAQDQOSIILVDSAILVRKE-----SHPLAFECOT 185  
 DB 385 KKMHEMLNPFDDSDKODTAFNELKKEITAHSAANLYGLNAPLITSLIADPRANEIOA 444  
 QY 186 LDNNEIVATIKFTVYTSELQWRSSLPATDAALIFVLTIGVLCVFIIFLIIPIIWA 245  
 DB 445 MEIRIPRAGPAVASEILMNRHILSPTKILLPFTTPII-----S 488  
 QY 246 AVKAFWAKASTPE-VQSEQSSVRK-----DSTSLDQLTTEPGE---DDALSE 291  
 DB 489 TIAPFKRKKTPQVQPEKKEPYKQRLKLDAAAKKISAFIPQGMTEQALKNIDE 548  
 QY 292 WNE 294  
 DB 549 WNO 551

## RESULT 12

Q9FCY5 PRELIMINARY; PRT; 456 AA.  
 AC Q9FCY5;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE Invasion protein IbeA.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 NC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96029746; PubMed=7591087;  
 RA Huang S.H., Wass C., Pu Q., Prasad Rao N.V., Stins M., Kim K.S.;  
 RT "Escherichia coli invasion of brain microvascular endothelial cells in  
 RT vitro and in vivo: molecular cloning and characterization of invasion  
 RT gene IbeA.";  
 RL Infect. Immun. 63:4470-4475 (1995).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20465456; PubMed=11008113; DOI=10.1016/S1286-4579(00)01277-6;  
 RA Huang S.H., Stins M.P., Kim K.S.;  
 RT "Bacterial penetration across the blood-brain barrier during the  
 RT development of neonatal meningitis.";  
 RL Microbes Infect. 2:1237-1244 (2000).

RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21136444; PubMed=11237832;  
 RA Huang S.H., Wan Z.S., Chen Y.H., Jong A.Y., Kim K.S.;  
 RT "Further characterization of Escherichia coli brain microvascular  
 RT endothelial cell invasion gene IbeA by deletion, complementation, and  
 RT protein expression.";  
 RL J. Infect. Dis. 183:1071-1078 (2001).

RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21652705; PubMed=11793250;  
 RA Huang S.H., Chen Y.H., Kong G., Chen S.H.M., Besemer J.,  
 RA Borodovsky M., Jong A.;  
 RT "A novel genetic island of meningitic Escherichia coli K1 containing  
 RT the IbeA invasion gene (IbeA): functional annotation and carbon-  
 RT source-regulated invasion of human brain microvascular endothelial  
 RT cells.";  
 RL Funct. Integr. Genomics 1:312-322 (2001).  
 DR EMBL; AF289032; AAP8391.2; --.  
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000205; NBD\_B8.  
 DR InterPro; IPR000103; Pyridine\_redox\_2.  
 DR PRINTS; PR00469; PNDRODTASRII.  
 SQ SEQUENCE 456 AA; 49772 MW; 00FD72D79B0B010 CRC64;

Query Match 6.4%; Score 97.5; DB 2; Length 456;  
 Best Local Similarity 18.9%; Pred. No. 9.2;  
 Matches 53; Conservative 39; Mismatches 96; Indels 93; Gaps 10;

QY 7 GCSAGLMTVGMILLAGLSARGTNVTAQDAGLAHEGSEETENNDSSTANYPPE 66  
 DB 49 GCFGMMWT-----AGESTAMRHENTVASSGLAIRB-----ETASSWGASS 92  
 QY 67 TEDVSNRVNVEVEFGMC--IVTCGIGVREV-----ILTNGCGESK 107  
 DB 93 PEPQNSGAINIEEFKLVADMLAQAGRVRLHTIYAVDIQGNLLGVITBSKSGRAI 152  
 QY 108 CVRVERBCRGFTDCGW--GKPISESLSVRLAC----- 138  
 DB 153 LANYIIDCTGADIAWFGAPFIR--EREELMCWTVPFS CANINKANFQONINSTPEYKG 211  
 QY 139 -----IHNSPLARFK-YMKKLLAQDQOSIILVND-----SAILVRKES 176  
 DB 212 DWGADENKMKSYVHSCRMFSPYLAKVLRKRSAGIIFKDYTLTGGSMSTVTEYGDAN 271  
 QY 177 HPLAFECDTLDNNEIVATIKFTVYTSELQWRSSLPATDA 217  
 DB 272 YLVNVSIPAVDCTVFD-----LTRALIBGRKQAMQAI 306

## RESULT 13

KRP1\_SCHPO

ID KRL1\_SCHPO STANDARD; PRT; 709 AA.  
AC Q09175;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Dibasic processing endoprotease precursor (EC 3.4.21.-) (KEX2-related protease)  
DN Name=krlp1; Synonyms=krlp; ORFNames=SPAC22E12.09c;  
OS Schizosaccharomyces pombe (Fission yeast);  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
CC Schizosaccharomycetes; Schizosaccharomycetaceae;  
OX NCBI\_TaxId=4896;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=EG545;  
RX MEDLINE=95112801; PubMed=7813430;  
RA Davey J., Davis K., Imai Y., Yamamoto M., Matthews G.;  
RT "Isolation and characterization of Krp, a dibasic endopeptidase required for cell viability in the fission yeast Schizosaccharomyces pombe.";  
RT EMBO J. 13:5910-5921(1994).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooke K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K., James K., Jones L., Jones M., Leach S., McDonald S., McLean C., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Voiclaert G., Aert R., Robben J., Grymoprez B., Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu B., Dreano S., Gloux S., Lelaure V., Motier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huret S.M., Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerutti L., Lowe T., McCombie W.R., Paulsen O., Potholkin J., Sipakovski G.V., Useary D., Barrett B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RT Nature 415:871-880(2002).  
RL -FUNCTION: Membrane-bound, subtilisin-like serine protease that processes the P-factor precursor and other precursor proteins. Essential for cell viability.  
CC -CATALYTIC ACTIVITY: Cleaves substrate on the C-terminal side of dibasic residues.  
CC -COFACTOR: Calcium.  
CC -SUBCELLULAR LOCATION: Type I membrane protein. Late Golgi compartment (By similarity).  
CC -PTM: N-glycosylated.  
CC -SIMILARITY: Belongs to the peptidase S8 family. Furin subfamily.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X83435; CAAS7818.1; -;  
CC EMBL: Z70043; CAA93896.1; -;

DR PIR; S51793; S51793.  
DR HSSP; P13134; 10T5.  
DR MEROPS; S08.070; -;  
DR GenDB\_SpPombe; SPAC22E12.09c; -;  
DR InterPro; IPR000209; Pept\_S8\_S53.  
DR InterPro; IPR002884; Eprotonconvertap.  
DR Pfam; PF01483; E\_protein; 1.  
DR Pfam; PF00082; Peptidase\_S8; 1.  
DR PRINTS; PR00723; SUBTILISIN.  
DR ProDom; PD000717; Eprotonconvertap; 1.  
DR PROSITE; PS00136; SUBTILASE ASP; 1.  
DR PROSITE; PS00137; SUBTILASE HIS; 1.  
DR PROSITE; PS00138; SUBTILASE SER; 1.  
DR KX Calcium; Glycoprotein; Hydrolyase; Serine protease; Signal;  
RV Transmembrane; Zymogen.  
FT SIGNAL 1 22 Potential.  
FT PROPEP 23 82 Potential.  
FT PROPEP 83 102 Potential.  
FT CHAIN 103 709 Potential.  
FT DOMAIN 103 668 Lumenal (Potential).  
FT TRANSMEM 669 693 Potential.  
FT DOMAIN 694 709 Cytoplasmic (Potential).  
FT DOMAIN 612 647 Ser/Thr-rich.  
FT ACT\_SITE 162 162 Charge relay system (By similarity).  
FT ACT\_SITE 200 200 Charge relay system (By similarity).  
FT ACT\_SITE 371 371 Charge relay system (By similarity).  
FT DISULFD 216 363 By similarity.  
FT CARBOHYD 308 338 By similarity.  
FT CARBOHYD 155 155 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 463 463 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 471 471 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 620 620 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 709 AA; 78126 MW; 414FE9B89CBE0840 CRC64;  
Query March 6.3%; Score 97; DB 1; Length 709;  
Best local similarity 25.8%; Pred. No. 17;  
Matches 40; Conservative 17; Mismatches 52; Indels 46; Gaps 7;  
QY 150 MWKLLRDOOQ-----SILVNDASILEVKESHPLAFECPTLD----- 187  
DB 558 VMKLVNDRSGKKEGTFENMQGLAWGDS---ENPSTAPLPY--DTLELPKENVLGIYS 612  
QY 188 --NNEIV-----ATIKTQVTVSSSLQMRSS-----LPATDAAL-IFVLTIGVI 228  
DB 613 EPNEDLTNNSSTLSPSTSTSYVSATAPTSHTPIPTVLPPTQVLEPSYREIVAF 672  
QY 229 ICFEILFLPIILINMAVKAFAKASTPEVQSE 263  
DB 673 ITFPLFAFLFVAVITWISAFWAKAPPLSQQE 707  
RESULT 14  
ID Q9V9G9 PRELIMINARY; PRT; 1014 AA.  
AC Q9V9G9; Q9SR33;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 25-OCT-2004 (TREMBLrel. 26, Last annotation update)  
DE Cg31619-PB (CHI19218p)  
DN Name=Cg3131; ORFNames=Cg31619;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
CC Ephydroidae; Drosophilidae; Drosophila.  
OX NCBI\_TaxId=7227;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celitker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,



RN [2]

**RN [3]**

RN [4]

RN [5]

1. The first group of people who are not in the labor force are those who are not in the labor force because they are not in the labor force.

RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abell J.F., Agbayant A., An H.J., Andrews-Framkoch C., Baldwin D.,  
 RA Ballew R.M., Baas A., Bakendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhander D., Bolshakov S.,  
 RA Borckova D., Botchan M.R., Bouck J., Brockett P., Broctier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegawa C.,  
 RA Jajall M., Kalush F., Karpén G.H., Ke Z., Kennison J.A., Kethum K.A.,  
 RA Kimmel B.B., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mikhlov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,  
 RA Palazolo M., Pletman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler P., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of Drosophila melanogaster.",  
 RA Science 287:2185-2195(2000).  
 [2]  
 RA SEQUENCE FROM N.A.  
 RA MEDLINE=22426065; PubMed=12537568;  
 RA Cealiker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Fribe E., Hodgson A.,  
 RA George R.A., Hoekins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren B.J.,  
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.B., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RA "Finishing a whole-genome shotgun. Release 3 of the Drosophila  
 RA melanogaster euchromatic genome sequence.",  
 RA Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 [3]  
 RA SEQUENCE FROM N.A.  
 RA MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Fribe E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Cealiker S.E.,  
 RA "The transposable elements of the Drosophila melanogaster euchromatin:  
 RA a genomic perspective.",  
 RA Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 [4]  
 RA SEQUENCE FROM N.A.  
 RA MEDLINE=22426069; PubMed=12537572;  
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu L., Bernan B.P.,  
 RA Bertencourt B.R., Cealiker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.,  
 RA "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RA systematic review.",  
 RA Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 [5]  
 RA SEQUENCE FROM N.A.  
 RA FlyBase;  
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 [6]

RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003781; AA057229.2;  
 DR FlyBase; FBgn0051619; CG31619.  
 DR InterPro; IPR007110; IG-1like.  
 DR InterPro; IPR003598; IG-C2.  
 DR InterPro; IPR010909; PLAC.  
 DR InterPro; IPR000884; TSPL.  
 DR Pfam; PF00047; IG\_1.  
 DR Pfam; PF00090; TSP\_1; 8.  
 DR SMART; SM00408; IG2; 1.  
 DR SMART; SM00209; TSP1; 11.  
 DR PROSITE; PS0835; IG LIKE; 1.  
 DR PROSITE; PS0900; PLAC; 1.  
 DR PROSITE; PS0092; TSP1; 8.  
 SQ SEQUENCE 1353 AA; 150210 MW; F83CDB090964272P CRC64;  
 Query Match 6.2%; Score 95; DB 2; Length 1353;  
 Best Local Similarity 21.9%; Pred. No. 55;  
 Matches 61; Conservative 33; Mismatches 88; Indels 96; Gaps 14;  
 QY 8 CSAGL-----LMTVGMILLAGLQSGAGTNTVAVQ---DAGLAHGGGEETENNDSSET 58  
 DB 419 CSTGIHCGSLNKKVGGTIIVG--SSRSINRSERQLOSDADENEDENEDGDDVDLES 476  
 QY 59 AEN-----YA-----PPEIVEDVNRN----- 74  
 DB 477 GQDDDDGEGSYADQPLLYAHTQSRLNQAAPDEPRTHLWNGSNPNFNGEDESSEGPS 536  
 QY 75 -----VXKEVERGMCTVTCGIVREVLITNGCP--GGSEKCVRYEE--CRGPTDGMGKP 126  
 DB 537 LDPIYIKDNENSPSCVTCGSEIRK--RTYCKITLFLYSRTVATYNDLSCEBK-----XP 588  
 QY 127 ISESLSEYRLACIHTS-----PLNRFK-----YMKKLARDQDSIILVNDAIL 170  
 DB 589 HDEVERCYEDCPMLPSHGPDQPPRDSIKGVSEPGKTYVVR-----EGYTSASAGLIG 643  
 QY 171 EVKRESHPLEACDPTDNNELVATIKFTVYTSSELQNR 208  
 DB 644 GVER-----LIINCVRBNDRGVSPFLCSPFTKPARVR 677

Search completed: April 15, 2005, 14:15:15  
 Job time : 192 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2005, 14:08:19 ; Search time 22 Seconds  
(without alignments)  
997,583 Million cell updates/sec

Title: US-10-809-655-9

Perfect score: 1528  
Sequence: 1 MSFRTGCGAGLMTVGMILL.....LQPLTFMGEDDLSERN 294

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgml2\_6/prodata/1/1aa/5A\_COMB.pep.\*  
2: /cgml2\_6/prodata/1/1aa/5B\_COMB.pep.\*  
3: /cgml2\_6/prodata/1/1aa/6A\_COMB.pep.\*  
4: /cgml2\_6/prodata/1/1aa/6B\_COMB.pep.\*  
5: /cgml2\_6/prodata/1/1aa/6CTUS\_COMB.pep.\*  
6: /cgml2\_6/prodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	106	6.9	479	4 US-09-270-767-46823	Sequence 46823, A
2	89.5	5.9	2181	4 US-09-949-016-5981	Sequence 5981, Ap
3	89.5	5.9	2188	4 US-09-949-016-8295	Sequence 8295, Ap
4	85	5.6	467	3 US-09-046-736-2	Sequence 2, Appl
5	85	5.6	596	4 US-09-949-016-7821	Sequence 7821, Ap
6	83.5	5.5	301	5 PCT-US95-13975-72	Sequence 72, Appl
7	83.5	5.5	618	4 US-08-311-731A-158	Sequence 158, App
8	83	5.4	519	4 US-09-248-796A-14534	Sequence 14534, A
9	82.5	5.4	366	3 US-09-134-001C-4799	Sequence 4799, Ap
10	80	5.2	566	4 US-09-513-783A-142	Sequence 142, App
11	80	5.2	566	4 US-09-430-656-142	Sequence 142, App
12	80	5.2	794	1 US-07-885-972A-2	Sequence 2, Appl
13	80	5.2	794	1 US-07-885-972A-4	Sequence 4, Appl
14	80	5.2	794	1 US-08-865-203-2	Sequence 2, Appl
15	80	5.2	794	2 US-08-745-880-2	Sequence 2, Appl
16	80	5.2	794	2 US-08-745-880-4	Sequence 4, Appl
17	80	5.2	794	2 US-08-480-382-2	Sequence 2, Appl
18	80	5.2	794	2 US-08-480-382-4	Sequence 4, Appl
19	80	5.2	794	2 US-07-849-420-2	Sequence 2, Appl
20	80	5.2	794	3 US-09-253-854-2	Sequence 2, Appl
21	80	5.2	794	3 US-08-955-424-2	Sequence 2, Appl
22	80	5.2	794	4 US-09-592-480-1	Sequence 1, Appl
23	80	5.2	794	4 US-10-133-910-2	Sequence 2, Appl
24	80	5.2	794	4 US-09-949-016-6239	Sequence 6239, Ap
25	80	5.2	812	4 US-09-513-783A-4	Sequence 4, Appl
26	80	5.2	812	4 US-09-513-783A-6	Sequence 6, Appl
27	80	5.2	812	4 US-09-430-656-4	Sequence 4, Appl

28	80	5.2	812	4 US-09-430-656-6	Sequence 6, Appl
29	80	5.2	820	4 US-09-949-016-9924	Sequence 9924, Ap
30	80	5.2	1056	4 US-09-513-783A-32	Sequence 32, Appl
31	80	5.2	1056	4 US-09-430-656-32	Sequence 32, Appl
32	80	5.2	1125	4 US-09-513-783A-152	Sequence 152, App
33	80	5.2	1125	4 US-09-430-656-152	Sequence 152, App
34	80	5.2	1610	4 US-09-513-783A-22	Sequence 22, Appl
35	80	5.2	1610	4 US-09-430-656-22	Sequence 22, Appl
36	80	5.2	1658	4 US-09-902-540-12643	Sequence 12643, A
37	79	5.2	582	3 US-09-419-459-2	Sequence 2, Appl
38	78.5	5.1	546	3 US-08-935-855-20	Sequence 20, Appl
39	78.5	5.1	546	4 US-09-538-092-827	Sequence 827, App
40	78.5	5.1	546	4 US-09-949-016-6265	Sequence 6265, App
41	78.5	5.1	553	4 US-09-949-016-7397	Sequence 7397, Ap
42	78	5.1	546	4 US-09-540-236-2574	Sequence 2574, Ap
43	78	5.1	935	1 US-07-707-367-2	Sequence 2, Appl
44	78	5.1	1096	4 US-09-540-236-2620	Sequence 2620, Ap
45	77.5	5.1	381	4 US-09-673-395A-363	Sequence 363, App

ALIGNMENTS

```
RESULT 1
US-09-270-767-46823
; Sequence 46823, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 46823
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46823

Query Match          6.9%; Score 106; DB 4; Length 479;
Best Local Similarity 22.1%; Pred. No. 0.0085;
Matches 62; Conservative 36; Mismatches 102; Indels 80; Gaps 12;

QY 43 HEGGEGEETENDSETAENYAPPEDEVSNRVKVEFG-----MCTVCGIGVR 93
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 204 HGRGSGQ-----SRVCMHACPAEQLSSNDNEHGEWCMGEMGACSVTCGLGR 257
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 94 EVILNCGCPGQ-----ESKC-VYVRECRGPTDGMKPISELSRYLACIHT 141
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 258 R-RTTRCIAGHDRLCGRALBQKCEWPCEDFLGWS--ANSEWSSCSGDIRL----- 308
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 142 SPLNRKYMVKLLRQDQSIILVNDASILEVRKSHPLAFECDTLDNNEIVATIKFTVYT 201
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 309 -----RHRCLVEQ-----FGSMKRG-----APFRTACV 334
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 202 SSELQWRSSLPATDAALFVLTIGVICVFIIILI---FIILNAAVAFGAAATSP 258
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 335 PNECECTQASATLPIVIFVGLFTVACCLATYRTKRFMLASBAELINKTTTASFD 394
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 259 EVQSEGSVRYDSTSLDLPF-----EMPGEDDLSERN 293
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 395 TYFNQYSLSLPTD--YYDQRPKRQSSFRMPATSNLGN 432
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 2
US-09-949-016-5981
; Sequence 5981, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5981  
LENGTH: 2181  
TYPE: PRF  
ORGANISM: Human  
US-09-949-016-8295

Query Match 5.9%; Score 89.5; DB 4; Length 2181;  
Best Local Similarity 21.2%; Pred. No. 6.1;  
Matches 71; Conservative 46; Mismatches 97; Indels 121; Gaps 18;

QY 32 VTAAVADAGLAHGEGBEETENDSETAENYAPPTEDVSNRVVKEVFGMCTTCGIG 91  
DB 442 ITQA-EDIPENBEBGEGKNTSMPTS-----ETESVTENVSGEGR----- 484  
QY 92 VAEVILNCGCPG-----GSKCVVVEBCRGPTDC-GMGKPISESLSVR----- 135  
DB 485 -----NRCCGSLCWMRRRGAAGA-----GSGCRRWGQALSKSLRRRRMRNR 530  
QY 136 -----LACHTSPLNKPKYM-----WKLRD-OOSIILVNDALIL 170  
DB 531 FNRRCRAAVKSVTFYVLYIVLPLNTLTLSSEHNQPMLOIDIANKVLLALPTCEM 590  
QY 171 EVRKESHP-----AFEC-----DTLNNEIVATIKFTVTSSEL-----QNR 208  
DB 591 LVKMTSLGQAYFVSLFNRFPDCEVCGITETILVELEIMSPGISVFCVRLRIFKVT 650  
QY 209 R--SSLPATDALI-FVLITGVILCFIIFLFIILINMAAVKAFGAKASTPEVQSEOS 265  
DB 651 RMTSLSNLWASLMSKMSIASLILLLPLFIIFSLG-----MOLFQKFNDEYQTKS 706  
QY 266 SVRYKSTSLDLPTE-----MPGEDDALSEMN 293  
DB 707 -----TFDNFPQALLTVFQILTGED-----WN 728

RESULT 3  
US-09-949-016-8295  
Sequence 8295, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8295  
LENGTH: 2188  
TYPE: PRF  
ORGANISM: Human  
US-09-949-016-8295

Query Match 5.9%; Score 89.5; DB 4; Length 2188;

Best Local Similarity 21.2%; Pred. No. 6.1;  
Matches 71; Conservative 46; Mismatches 97; Indels 121; Gaps 18;

QY 32 VTAAVADAGLAHGEGBEETENDSETAENYAPPTEDVSNRVVKEVFGMCTTCGIG 91  
DB 449 ITQA-EDIPENBEBGEGKNTSMPTS-----ETESVTENVSGEGR----- 491  
QY 92 VAEVILNCGCPG-----GSKCVVVEBCRGPTDC-GMGKPISESLSVR----- 135  
DB 492 -----NRCCGSLCWMRRRGAAGA-----GSGCRRWGQALSKSLRRRRMRNR 537  
QY 136 -----LACHTSPLNKPKYM-----WKLRD-OOSIILVNDALIL 170  
DB 538 FNRRCRAAVKSVTFYVLYIVLPLNTLTLSSEHNQPMLOIDIANKVLLALPTCEM 597  
QY 171 EVRKESHP-----AFEC-----DTLNNEIVATIKFTVTSSEL-----QNR 208  
DB 598 LVKMTSLGQAYFVSLFNRFPDCEVCGITETILVELEIMSPGISVFCVRLRIFKVT 657  
QY 209 R--SSLPATDALI-FVLITGVILCFIIFLFIILINMAAVKAFGAKASTPEVQSEOS 265  
DB 658 RMTSLSNLWASLMSKMSIASLILLLPLFIIFSLG-----MOLFQKFNDEYQTKS 713  
QY 266 SVRYKSTSLDLPTE-----MPGEDDALSEMN 293  
DB 714 -----TFDNFPQALLTVFQILTGED-----WN 735

RESULT 4  
US-09-046-736-2  
Sequence 2, Application US/09046736  
Patent No. 6090582  
GENERAL INFORMATION:  
APPLICANT: KITKY, KRISTINE  
TITLE OF INVENTION: ERICKSON-WILDER, CONNIE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,736  
FILING DATE: 24-MAR-1998  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/041,885  
FILING DATE: 02-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-50019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-046-736-2

```

Query Match 5.6%; Score 85; DB 3; Length 467;
Best Local Similarity 19.8%; Pred. No. 1.8; Indels 84; Gaps 11;
Matches 59; Conservative 39; Mismatches 116; Indels 84; Gaps 11;

QY 6 TCGSAGLLMTVGMLLAG---LQSRGINTVTAVADAGLAHGBGEBEETNDSETAENV 62
DB 160 SGCFFNLCSVPMACEQGTTPMISMMGTSP----- 191
QY 63 APPEFEDVSNRNVAKEVEFGMCTVTC-----GIV---REVITTCGPGESKCVRYVE 114
DB 192 PHSPTRSSVLTLLIPQPHHGTSLTCQVTLPGAGVYTNRTIQLNVSP--PQNLTTVTPQ 249
QY 115 CRGPTDCMGKRISSSL---ESVPLAC-IHTSPNPKMMKLLRQDOSIILVNDALI 170
DB 250 GBGTAISTALGNSSSLVLEGOSLRVCADVSDNPALISWTR-----SLTL 295
QY 171 EVRKESHPLAFECDTLDNNEIVATTIKETVYTSSELQMRSSL-----PATDA 217
DB 296 YPSQSNLVLVELQVHLGDE---GFTCRAGNSLOSCHVSLSLQGYTGMRVPSGV 351
QY 218 ALIFVLTGVIIICVFIIIFLIIIIINMAAVKAFWAKASTPEVQSSQSVRYDOSTSL 275
DB 352 LLGAVGAGATATVLFSCVFIEIVR-----SCRKTSAPADVDGVGMKDANTI 401

RESULT 5
US-09-949-016-7821
; Sequence 7821, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001037
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7821
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7821

Query Match 5.6%; Score 85; DB 4; Length 596;
Best Local Similarity 25.7%; Pred. No. 2.6;
Matches 43; Conservative 23; Mismatches 55; Indels 46; Gaps 9;

QY 46 EGEETEND-SETAENVAPPEFEDVSNRNVAKEVEFGMCTVTCGIVREVITTCGPG 104
DB 229 BGPATVSSKASDITVEGVLPSSKQEPENDGVVSQ-----AGXDRSKLLE-----G 274
QY 105 ESKCVARYEBCGPTDCMGKRPISSELSVRLACIHTSPNPKVM-----W----- 151
DB 275 KAADIKOBE---PGDLGRPKCEBSYDIPNALYCIRQPHNN-RFNICCDRCGEWTFHDC 350
QY 152 -----KLARQDOSIILVNDALIEVRKESHPLAFECDTLDNNE 190
DB 331 VGISARGRLLERNGEDYICPN-CTILOVODETH-----SETNDQOE 371

RESULT 6
PCT-US95-13975-72
; Sequence 72, Application PC/TUS9513975
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: McDowell, Michael W.

```

```

TITLE OF INVENTION: Recombinant Feline Herpes virus
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSER: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 310 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13975
FILING DATE: 26-OCT-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/329,883
FILING DATE: 26-OCT-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39118-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212)391-0525
INFORMATION FOR SEQ. ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 301 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
PCT-US95-13975-72

Query Match 5.5%; Score 83.5; DB 5; Length 301;
Best Local Similarity 19.8%; Pred. No. 1.3;
Matches 59; Conservative 43; Mismatches 99; Indels 97; Gaps 12;

QY 4 RGNCSAGILMTVGMILLAGLSARGNVTAAY---QDAGLAHRCGEETENNDSFAE 60
DB 81 RGRCCPEPELVEPLVMV-----NISDDLGCFFPPYHNHEDRDISDELPEEDIH 129
QY 61 NYAPPEDEVSNENVKVEFGMCVTGIGVREVLITNGCPGGSXCVRVVECRGPTD 120
DB 130 NYS-----DCRATNM 139
QY 121 CGMGKPISELSVRLAC---IHTSPLNRFKMKKLIRDOQSITLVNDSAILVEYR 174
DB 140 FVPRBPLSQVLGQSOLVGSIGRQIITQP-----WQ-LKQNSYDGLRV--ASLEPRH 189
QY 175 ESHPLAECCDTLNNELIVATIKFTVYTSSELOWRBSL-----PATDALIF--VLTI 225
DB 190 LDS--SNDRLDLDETEGSIITPPETHKGVNGGFLDPLIIEPTPELVHTKIKGI 247
QY 226 GVLIICVFIILFIILINMAVAFAFWGAKASTEVESSVRYKSTSLDQLEPTMP 283
DB 248 GTVVVVFLLIILSLCYVTCVLSRIGMD-----RAYVQGVFNHNSPSYQL-TRYP 299

RESULT 7
US-08-311-731A-158
Sequence 158, Application US/08311731A
Patent No. 6583266
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
NUMBER OF SEQUENCES: 411

```

CORRESPONDENCE ADDRESS:  
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
STREET: 600 ATLANTIC AVENUE  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311,731A  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: GATES, EDWARD R.  
REGISTRATION NUMBER: 31,616  
REFERENCE/DOCKET NUMBER: C0044/7125  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/720-3500  
TELEFAX: 617/720-2441  
INFORMATION FOR SEQ ID NO: 158:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 618 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Mycobacterium leprae  
US-08-311-731A-158

Query Match 5.5%; Score 83.5; DB 4; Length 618;  
Best Local Similarity 24.1%; Pred. No. 4;  
Matches 60; Conservative 30; Mismatches 102; Indels 57; Gaps 14;

QY 4 RTGCGAGLMTVGMWTL--LLAGLQASRG-TNTTAADVQAGLAHGBGEE-ETENNDSST 58  
DB 58 RALANQAGVGTSGMKRSELIAIEECRGANGTSVNDGSPSRDHGSATNISTBALAQS 117  
QY 59 AANNAPPEDEDVSNRVVVEVERGMCVTVCIGIVREVIL-----TNGCGGSEKC 108  
DB 118 EONYALVEVSRERRRARSAD-----VTAISTATETSDCCGTADDTRTIOGGQSD- 171  
QY 109 VVREBCRGPTDCGKGPISSESLSVRLACIHTSP-----NRFKYMKLAR 155  
DB 172 -KTER-RKP-DVGNQGVTEQSSSLQ-----PRGDDDEGRGGRGRFRDRRRR 221  
QY 156 QDOQSILVNDAILVERKES--HPLAFECDTLDNREIVATIF-----TVTSSSELQMR 208  
DB 222 GERS-----GDGARELRQDDVQVQVAGILDVLDNVAFTVTSGLAGPHDVVYSMSM-VR 275  
QY 209 RSLSPATDA 217  
DB 276 KNGLRGRDA 284

RESULT 8  
US-09-248-796A-14534  
Sequence 14534, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Ketch Wainstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 14534  
LENGTH: 519  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-14534

Query Match 5.4%; Score 83; DB 4; Length 519;  
Best Local Similarity 22.9%; Pred. No. 3.5;  
Matches 30; Conservative 30; Mismatches 47; Indels 24; Gaps 5;

QY 127 ISSELSVRLACIHTSPPLRPFYKMKLIRDOQSIIIVNDAILVERKESHPLAFECDTL 186  
DB 25 LAETLEETALIIISVLSFINGRSH-----KQEEESTLSTNNNN-----NNDNN 66  
QY 187 DNEEIVATIKFTVYTSSELQMRSSLPATDAALIFV-LTIGVITC--VFITLLFIITN 243  
DB 67 DNNRIENENNAVHPDISISQSRKADNRRKLKPOVIGAILGIIIFVLIFFVFG 126  
QY 244 ---VAAVRAFW 251  
DB 127 ODVWSYTERVW 137

RESULT 9  
US-09-134-001C-4799  
Sequence 4799, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4799

LENGTH: 366  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4799

Query Match 5.4%; Score 82.5; DB 3; Length 366;  
Best Local Similarity 19.0%; Pred. No. 2.3;  
Matches 47; Conservative 41; Mismatches 80; Indels 79; Gaps 11;

QY 21 LAGLQASRGNTNTTAADVQAGLAHGBGEEETENNDS-----ETANNVAPPEDEDVSN 72  
DB 107 IPGISGVDRSITRKIRNYGVLRAG---FTDNKNDIOELVQLKTAB-LPRDEVQTVST 161  
QY 73 R-----NVKEVEFGMCTVT-----CGIGVREVILITNG 100  
DB 162 KPIPVYSTGSDLSVLLDPGKQNIIVRELNRGCVTVVPRDYSABEILGMSPGVWLSNG 221  
QY 101 CPGESEKCVVREBCRGPTDCGK-----DISSELSVRLACIHTSPPLNR 147  
DB 222 -PEDPDEVVALDMIRGIL-----GKIPFGICLGHQLPALSQATSFKMKFGRGCA----- 272  
QY 148 KYMKKLIRDOQSIIIVNDAILVERKESHPLAFECDTLDNREIVATIFVTYTSSELQMR 207  
DB 273 NHEVQDLKFTGK-----IDITSQNHGYSIDCDLSKNTDLEVT-HIALNDGVTEGL 320  
QY 208 RSLSPRA 214  
DB 321 RHEKELRA 327

RESULT 10

US-09-513-783A-142  
Sequence 142, Application US/09513783A  
Patent No. 6416959  
GENERAL INFORMATION:  
APPLICANT: Giuliano, Kenneth A.  
APPLICANT: Kapur, Ravi  
TITLE OF INVENTION: A System for Cell Based Screening  
FILE REFERENCE: 97-022-11  
CURRENT APPLICATION NUMBER: US/09/513,783A  
CURRENT FILING DATE: 2000-02-25  
NUMBER OF SEQ ID NOS: 180  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 142  
LENGTH: 566  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Size exclusion  
US-09-513-783A-142

Query Match 5.2%; Score 80; DB 4; Length 566;  
Best Local Similarity 22.1%; Pred. No. 8.7;  
Matches 53; Conservative 32; Mismatches 101; Indels 54; Gaps 7;

QY 2 SPRGTCGSAGLMTVGMILLAGIQAAGTAVTAQVODAGLAHGEGBEETENNDSYAE 61  
DB 344 SPKETTTPKRD-----LAPPEVLLTKETELAPAKGVNSISEBALAKNDVSAEI 398  
QY 62 YAPPETEDVSNRVVVEFGMCT-----VTCGIGVREVLITNGCPGES 106  
DB 399 PVAQET-----VSETEVVLATVLPSPDPTTLTKOVTLPLEBARPLVTMTSLET 451  
QY 107 KCVAVVECGRPDCCGK-----PISLSLSEVRLACHTSPLNFKYMTMLLRODQ 159  
DB 452 EMTLG-KETAPPETETMGNAKDMSPLBESHVTLGKDVLLPEKVAEFNNVTPLSEBEVT 510  
QY 160 SILVNDASILEVRKSHPLAFECPTLDNNEIVATIKFTYTSSELOMRSSLPATDAAL 219  
DB 511 SVKMSPSAETEA-----PLAKNAD-----LHSGTELVDNSMAPASDLAL 551

RESULT 11  
US-09-430-656-142  
Sequence 142, Application US/09430656  
Patent No. 6756207  
GENERAL INFORMATION:  
APPLICANT: Giuliano, Kenneth A.  
APPLICANT: Bright, Gary  
APPLICANT: Olson, Keith  
APPLICANT: Burroughs-Tencza, Sarah  
TITLE OF INVENTION: A System for Cell Based Screening  
FILE REFERENCE: 97-022-K  
CURRENT APPLICATION NUMBER: US/09/430,656  
CURRENT FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 09/398,965  
PRIOR FILING DATE: 1999-09-17  
PRIOR APPLICATION NUMBER: 09/031,271  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 08/810,983  
PRIOR FILING DATE: 1997-02-27  
PRIOR APPLICATION NUMBER: 60/136,078  
PRIOR FILING DATE: 1999-05-26  
PRIOR APPLICATION NUMBER: 60/106,308  
PRIOR FILING DATE: 1998-10-30  
NUMBER OF SEQ ID NOS: 168  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 142  
LENGTH: 566  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Size exclusion

OTHER INFORMATION: target sequence  
US-09-430-656-142

Query Match 5.2%; Score 80; DB 4; Length 566;  
Best Local Similarity 22.1%; Pred. No. 8.7;  
Matches 53; Conservative 32; Mismatches 101; Indels 54; Gaps 7;

QY 2 SPRGTCGSAGLMTVGMILLAGIQAAGTAVTAQVODAGLAHGEGBEETENNDSYAE 61  
DB 344 SPKETTTPKRD-----LAPPEVLLTKETELAPAKGVNSISEBALAKNDVSAEI 398  
QY 62 YAPPETEDVSNRVVVEFGMCT-----VTCGIGVREVLITNGCPGES 106  
DB 399 PVAQET-----VSETEVVLATVLPSPDPTTLTKOVTLPLEBARPLVTMTSLET 451  
QY 107 KCVAVVECGRPDCCGK-----PISLSLSEVRLACHTSPLNFKYMTMLLRODQ 159  
DB 452 EMTLG-KETAPPETETMGNAKDMSPLBESHVTLGKDVLLPEKVAEFNNVTPLSEBEVT 510  
QY 160 SILVNDASILEVRKSHPLAFECPTLDNNEIVATIKFTYTSSELOMRSSLPATDAAL 219  
DB 511 SVKMSPSAETEA-----PLAKNAD-----LHSGTELVDNSMAPASDLAL 551

RESULT 12  
US-07-885-972A-2  
Sequence 2, Application US/07885972A  
Patent No. 5460950  
GENERAL INFORMATION:  
APPLICANT: Barr, Phillip J.  
APPLICANT: Brake, Anthony J.  
APPLICANT: Kaufman, Rhadai J.  
APPLICANT: Tekamp-Olson, Patricia  
APPLICANT: Waseley, Louise  
APPLICANT: Wong, Polly A.  
TITLE OF INVENTION: Expression of PACB in Host Cells and  
TITLE OF INVENTION: Methods of Use Thereof  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Howson & Howson  
STREET: Spring House Corporate Center, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/885,972A  
FILING DATE: 19920520  
CLASSIFICATION: 435  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/621,092  
FILING DATE: 26-NOV-1990  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,859  
FILING DATE: 29-NOV-1990  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/621,443  
FILING DATE: 28-NOV-1990  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/621,457  
FILING DATE: 30-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary B.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: G15181A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 794 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-885-972A-2

Query Match 5.2%; Score 80; DB 1; Length 794;  
Best Local Similarity 21.7%; Pred. No. 15;  
Matches 75; Conservative 27; Mismatches 111; Indels 132; Gaps 18;

1 MSPRGTGSGAGLMTVGMILLAGL---QSARGNTVTAAVQDAGLAHSGEETENNDS 57  
506 VSPWGTIRST-----LLAPRPHDYSADGFNDMAFTTHSWDEPSGEVLTIENTS 555  
58 TAENY-----APPE-----TEDVSNRVNKEVEFGM-----CTVTCG 89  
556 EANNVGTILKFTLVLYGTAPBGLPVPBESSGCKTITSSQACVCEGFSLHQKSCVQHCP 615  
90 IGR-----EVILTNGCPGSESKCVRVBECRGP--TDGCMGKPISESL 133  
616 PGFAPQVLDTHYSTENDVETIRASVCAPCHASCAT---CQGFALTDG--LSCPSHASLDP 670  
134 VRLACIHTSPILRPFKMKMLRDPQDSIIVNDSALLEVRSKSHPLAFPCDTLDNNEIYA 193  
671 VQTC-----SRQSS-----RSP-----QOQPP 692  
194 TIKFTVYTSSEIQ--MRRSLPATDAALIFVLITGVLCVFI--IFLLFIINMAAVKA 249  
693 RLPEVEAGRLRAGLLPSHLPEVAVGL-----SCAFVLVFTVTFVLQLRSGPS 743  
250 FWGASTPEVOSBOSVRYKOSTSLDQLETPMPEGDALSEWNE 294  
744 FRGVKYYT---MDRGLISYKG---LPPE-----AMQE 769

RESULT 13

US-07-885-972A-4  
Sequence 4, Application US/07885972A  
Patent No. 5460950

GENERAL INFORMATION:  
APPLICANT: Bart, Phillip J.  
APPLICANT: Brake, Anthony J.  
APPLICANT: Kaufman, Rudael J.  
APPLICANT: Tekamp-Olson, Patricia  
APPLICANT: Wasley, Louise  
APPLICANT: Wong, Polly A.  
TITLE OF INVENTION: Expression of PACE in Host Cells and  
TITLE OF INVENTION: Methods of Use Thereof  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Howson & Howson  
STREET: Spring House Corporate Center, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/885,972A  
FILING DATE: 19920520  
CLASSIFICATION: 435  
APPLICATION NUMBER: 435  
PRIOR APPLICATION NUMBER: US 07/621,092  
FILING DATE: 26-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,859  
FILING DATE: 29-NOV-1990

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/621,443  
FILING DATE: 29-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/621,457  
FILING DATE: 30-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: G15181A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 794 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-885-972A-4

Query Match 5.2%; Score 80; DB 1; Length 794;  
Best Local Similarity 21.7%; Pred. No. 15;  
Matches 75; Conservative 27; Mismatches 111; Indels 132; Gaps 18;

1 MSPRGTGSGAGLMTVGMILLAGL---QSARGNTVTAAVQDAGLAHSGEETENNDS 57  
506 VSPWGTIRST-----LLAPRPHDYSADGFNDMAFTTHSWDEPSGEVLTIENTS 555  
58 TAENY-----APPE-----TEDVSNRVNKEVEFGM-----CTVTCG 89  
556 EANNVGTILKFTLVLYGTAPBGLPVPBESSGCKTITSSQACVCEGFSLHQKSCVQHCP 615  
90 IGR-----EVILTNGCPGSESKCVRVBECRGP--TDGCMGKPISESL 133  
616 PGFAPQVLDTHYSTENDVETIRASVCAPCHASCAT---CQGFALTDG--LSCPSHASLDP 670  
134 VRLACIHTSPILRPFKMKMLRDPQDSIIVNDSALLEVRSKSHPLAFPCDTLDNNEIYA 193  
671 VQTC-----SRQSS-----RSP-----QOQPP 692  
194 TIKFTVYTSSEIQ--MRRSLPATDAALIFVLITGVLCVFI--IFLLFIINMAAVKA 249  
693 RLPEVEAGRLRAGLLPSHLPEVAVGL-----SCAFVLVFTVTFVLQLRSGPS 743  
250 FWGASTPEVOSBOSVRYKOSTSLDQLETPMPEGDALSEWNE 294  
744 FRGVKYYT---MDRGLISYKG---LPPE-----AMQE 769

RESULT 14

US-08-865-203-2  
Sequence 2, Application US/08865203  
Patent No. 5935815

GENERAL INFORMATION:  
APPLICANT: van de Ven, Willem Jan Marie  
APPLICANT: van den Ouweland, Anna Maria Wilhelmina  
APPLICANT: van Duljnoven, Johannes Lambertus Petrus  
APPLICANT: Robroek, Antonius Johannes Maria  
APPLICANT: Konig, Piet Nico Maria  
TITLE OF INVENTION: Pharmaceutical Composition Having An  
TITLE OF INVENTION: Endoproteolytic Activity; A Process for  
TITLE OF INVENTION: Endoproteolytically Processing (Precursor)  
TITLE OF INVENTION: Proteins And For The (Micro)Biological  
TITLE OF INVENTION: Production Of Proteins  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSER: HOFMANN & BARON, LLP  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11753



COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/865,203  
FILING DATE: 29-MAY-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Tran, Jessica H.  
REGISTRATION NUMBER: 40,846  
REFERENCE/DOCKET NUMBER: 294-41 DIV II  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 794 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-865-203-2

Query Match 5.2%; Score 80; DB 2; Length 794;  
Best Local Similarity 21.7%; Pred. No. 15;  
Matches 75; Conservative 27; Mismatches 111; Indels 132; Gaps 18;

1 MSPRTGCSAGLMTVGMILLAGL---OSARGTNTVAVODAGLAHGEGBETENNDS 57  
506 VSPMGTRST-----LLAARPHDYADGFNDMAFWTTHSWDEDPGEVWLEIENIS 555  
58 TAEVY-----APPE-----TEDVSNRNVVKVEFGM-----CTVTCG 89  
556 EANNYGTLTKFTLVLYGTAPBGLPVPPESSGCKTLTSSQACVCEBGSFLHOKSCVQHCP 615  
90 IGVY-----EVILTNGCPGSGSKCVAVVEECRGP--TDCGMGXPISSLES 133  
616 PGFAPQVLDTHYSTENDVETIRASVCAPCASCAT---CQGPALTDG-LSCPSHSLDP 670  
134 VALACHTSPILNRFKYMMLLRDQOOSIILVNDASILEVRKESHPLAFECDTLDNNEIYA 193  
671 VEQTC-----SRQSOS-----RESP-----CQQPP 692  
194 TIKFTVYTSSEIQ--MRSSLPATDAALIFVLTIGVICVF--IFLLIFIIINMAAVKA 249  
693 RLPPEVAGORLRAGLPSHLPEVAVGL-----SCAFIVLVFVTVLVQLRSRGS 743  
250 FPGAKASTPEVQSEOSVRYKSTSLDQPTMPGDDALSEWNE 294  
744 FPGVAVYV---MDRGLISYKG---LPPE-----AMQE 769

RESULT 15  
US-08-745-880-2  
Sequence 2, Application US/08745880  
Patent No. 5965425  
GENERAL INFORMATION:  
APPLICANT: Bart, Phillip J.  
APPLICANT: Brake, Anthony J.  
APPLICANT: Kaufman, Rudal J.  
APPLICANT: Tekamp-Olson, Patricia  
APPLICANT: Waseley, Louise  
APPLICANT: Wong, Polly A.  
TITLE OF INVENTION: Expression of PAGE in Host Cells and  
TITLE OF INVENTION: Methods of Use Thereof  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESSES:  
ADDRESSES: Howson & Howson  
STREET: Spring House Corporate Center, P.O. Box 457  
CITY: Spring House

STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/745,880  
FILING DATE: 08-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US 08/480,382  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 07/885,972  
FILING DATE: 20-MAY-1992  
APPLICATION NUMBER: US 07/621,092  
FILING DATE: 26-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/620,859  
FILING DATE: 29-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/621,443  
FILING DATE: 29-NOV-1990  
APPLICATION NUMBER: US 07/621,457  
FILING DATE: 30-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: G15181A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-9818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 794 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-745-880-2

Query Match 5.2%; Score 80; DB 2; Length 794;  
Best Local Similarity 21.7%; Pred. No. 15;  
Matches 75; Conservative 27; Mismatches 111; Indels 132; Gaps 18;

1 MSPRTGCSAGLMTVGMILLAGL---OSARGTNTVAVODAGLAHGEGBETENNDS 57  
506 VSPMGTRST-----LLAARPHDYADGFNDMAFWTTHSWDEDPGEVWLEIENIS 555  
58 TAEVY-----APPE-----TEDVSNRNVVKVEFGM-----CTVTCG 89  
556 EANNYGTLTKFTLVLYGTAPBGLPVPPESSGCKTLTSSQACVCEBGSFLHOKSCVQHCP 615  
90 IGVY-----EVILTNGCPGSGSKCVAVVEECRGP--TDCGMGXPISSLES 133  
616 PGFAPQVLDTHYSTENDVETIRASVCAPCASCAT---CQGPALTDG-LSCPSHSLDP 670  
134 VALACHTSPILNRFKYMMLLRDQOOSIILVNDASILEVRKESHPLAFECDTLDNNEIYA 193  
671 VEQTC-----SRQSOS-----RESP-----CQQPP 692  
194 TIKFTVYTSSEIQ--MRSSLPATDAALIFVLTIGVICVF--IFLLIFIIINMAAVKA 249  
693 RLPPEVAGORLRAGLPSHLPEVAVGL-----SCAFIVLVFVTVLVQLRSRGS 743  
250 FPGAKASTPEVQSEOSVRYKSTSLDQPTMPGDDALSEWNE 294  
744 FPGVAVYV---MDRGLISYKG---LPPE-----AMQE 769

Search completed: April 15, 2005, 14:16:29

Mon Apr 18 12:47:48 2005

Job time : 24 secs

us-10-809-655-9.rad

Page 8